Giovanni Gallavotti

The Elements of Mechanics

Ipparco Editore, 2007

Giovanni Gallavotti
Dipartimento di Fisica
Università di Roma "La Sapienza"
Pl. Moro 2
00185, Roma, Italy
e-mail: giovanni.gallavotti@roma1.infn.it
web: http://ipparco.roma1.infn.it

A Daniela per amore infinito

Giovanni Gallavotti
Dipartimento di Fisica
Università di Roma "La Sapienza"
Pl. Moro 2
00185, Roma, Italy
e-mail: giovanni.gallavotti@roma1.infn.it
web: http://ipparco.roma1.infn.it

Preface

Preface to the Second English edition (2007).©

This is *Version* **1.3***: June 15, 2010*

In 2007 I recovered the Copyright. This is a new version that follows closely the first edition by Springer-Verlag. I made very few changes. Among them the Gauss' method, already inserted in the second Italian edition, has been included here. Believing that my knowledge of the English language has improved since the late '970's I have changed some words and constructions.

This version has been reproduced electronically (from the first edition) and quite a few errors might have crept in; they are compensated by the corrections that I have been able to introduce. This version will be updated regularly and typos or errors found will be amended: it is therefore wise to wait sometime before printing the file; the versions will be updated and numbered. The ones labeled 2.* or higher will have been entirely proofread at least once.

As owner of the Copyright I leave this book on my website for free downloading and distribution. *Optionally* the colleagues who download the book could send me a one line message (saying "downloaded", at least): I will be grateful. Please signal any errors, or sources of unhappiness, you spot.

On the web site I also put the codes that generate the non trivial figures and which provide rough attempts at reproducing results whose originals are in the quoted literature. Discovering the phenomena was a remarkable achievement: but reproducing them, having learnt what to do from the original works, is not really difficult if a reasonably good computer is available.

Typeset with the public Springer-Latex macros.

Giovanni Gallavotti — Roma 18, August 2007

Preface to the first English edition

The word "elements" in the title of this book does not convey the implication that its contents are "elementary" in the sense of "easy": it mainly means that no prerequisites are required, with the exception of some basic background in classical physics and calculus.

It also signifies "devoted to the foundations". In fact, the arguments chosen are all very classical, and the formal or technical developments of this century are absent, as well as a detailed treatment of such problems as the theory of the planetary motions and other very concrete mechanical problems. This second meaning, however, is the result of the necessity of finishing this work in a reasonable amount of time rather than an a priori choice.

Therefore a detailed review of the "few" results of ergodic theory, of the "many" results of statistical mechanics, of the classical theory of fields (elasticity and waves), and of quantum mechanics are also totally absent; they could constitute the subject of two additional volumes on mechanics.

This book grew out of several courses on “Meccanica Razionale”, i.e., essentially, Theoretical Mechanics, which I gave at the University of Rome during the years 1975-1978.

The subjects cover a wide range. Chapter 2, for example, could be used in an undergraduate course by students who have had basic training in classical physics; Chapters 3 and 4 could be used in an advanced course; while Chapter 5 might interest students who wish to delve more deeply into the subject, and fit could be used in a graduate course.

My desire to write a self-contained book that gradually proceeds from the very simple problems on the qualitative theory of ordinary differential equations to the more modem theory of stability led me to include arguments of mathematical analysis, in order to avoid having to refer too much to existing textbooks (e.g., see the basic theory of the ordinary differential equations in §2.2-§2.6 or the Fourier analysis in §2.13, etc.).

I have inserted many exercises, problems, and complements which are meant to illustrate and expand the theory proposed in the text, both to avoid excessive size of the book and to help the student to learn how to solve theoretical problems by himself. In Chapters 2-4, I have marked with an asterisk the problems which should be developed with the help of a teacher; the difficulty of the exercises and problems grows steadily throughout the book, together with the conciseness of the discussion.

The exercises include some very concrete ones which sometimes require the help of a programmable computer and the knowledge of some physical data. An algorithm for the solution of differential equations and some data tables are in Appendix O and Appendix P, respectively.

The exercises, problems, and complements must be considered as an important part of the book, necessary to a complete understanding of the theory.

In some sense they are even more important than the propositions selected for the proofs, since they illustrate several aspects and several examples and counterexamples that emerge from the proofs or that are naturally associated with them.

I have separated the proofs from the text: this has been done to facilitate reading comprehension by those who wish to skip all the proofs without losing continuity. This is particularly true for the more mathematically oriented sections. Too often students tend to confuse the understanding of a mathematical proposition with the logical contortions needed to put it into an objective, written form. So, before studying the proof of a statement, the student should meditate on its meaning with the help (if necessary) of the observations that follow it, possibly trying to read also the text of the exercises and problems at the end of each section (particularly in studying Chapters 3-5).

The student should bear in mind that he will have understood a theorem only when it appears to be self-evident and as needing no proof at all (which means that its proof should be present in its entirety in his mind, obvious and natural in all its aspects and, if necessary, describable in all details). This level of understanding can be reached only slowly through an analysis of several exercises, problem, examples, and careful thought.

I have illustrated various problems of classical mechanics, guided by the desire to propose always the analysis of simple rather than general cases. I have carefully avoided formulating "optimal" results and, in particular, have always stressed (by using them almost exclusively) my sympathy for the only "functions" that bear this name with dignity, i.e., the C^∞-functions and the elementary theory of integration ("Riemann integration").

I have tried to deal only with concrete problems which could be "constructively" solved (i.e., involving estimates of quantities which could actually be computed, at least in principle) and I hope to have avoided indulging in purely speculative or mathematical considerations. I realize that I have not been entirely successful and I apologize to those readers who agree with this point of view without, at the same time, accepting mathematically non rigorous treatments.

Finally, let me comment on the conspicuous absence of the basic elements of the classical theory of fluids. The only excuse that I can offer, other than that of non pertinence (which might seem a pretext to many), is that, perhaps, the contents of this book (and of Chapter 5 in particular) may serve as an introduction to this fascinating topic of mathematical physics.

The final sections, §5.9-§5.12, may be of some interest also to non students since they provide a self-contained exposition of Arnold's version of the Kolmogorov-Arnold-Moser theorem.

This book is an almost faithful translation of the Italian edition, with the addition of many problems and §5.12 and with §5.5, §5.7, and §5.12 rewritten.

I wish to thank my colleagues who helped me in the revision of the manuscript and I am indebted to Professor V. Franceschini for providing (from his files) the very nice graphs of §5.8.

I am grateful to Professor Luigi Radicati for the interest he showed in inviting me to write this book and providing the financial help from the Italian printer P. Boringhieri.

The English translation of this work was partially supported by the "Stiftung Volkswagenwerk" through the IHES.

Giovanni Gallavotti
Roma, 27 December 1981

Contents

1

Phenomena Reality and models

1.1 Statements

The results of physical experiments are determined by observations based on the measurement of various entities, i.e. the association of well defined sequences of numbers with well defined sequences of events.

The physical entities are "operationally defined". This means that they are defined in terms of the operations used to construct the numbers that provide their "measure".

For instance, the sequence of operations necessary to measure the "distance" between two given points P and Q in space consists in choosing a particular ruler and placing it on the straight line joining points P and Q, starting from P. Taking the endpoint of the ruler as the new starting point, the procedure is repeated n times until the endpoint of the ruler is superimposed on Q. If the distance PQ is not an exact multiple of the length of the ruler, one may, after n such operations, reach a point $Q_n \neq Q$ preceding Q on the line PQ; and after $n+1$ operations one may reach point Q_{n+1} following Q on the line PQ. Then one takes a new ruler "ten times shorter" and puts it on Q_nQ trying to match, as before, the second endpoint with Q. When this turns out to be impossible, one can, as in the first case, define a new point Q_{n_1} on Q_nQ and, then, take a third ruler ten times shorter than the second and repeat the operation.

Thus, inductively, a number $n + 0.n_1n_2\ldots$ (in decimal representation) is built which, by definition, is the measure of the distance between P and Q.

The above sequence of operations appears well defined but, in fact, a careful analysis shows that it does not have the prerequisites to be considered a

mathematically precise definition. What, for instance is "space", what is a "point", what is a "ruler"? Is it possible to "divide" a ruler into parts, and infinitely often?

The physicist is not too concerned (or, rather, not at all concerned) with such aspects of the question: he considers a physical entity well defined whenever the empirical procedure necessary for its measurement is clear.

A measurement procedure is considered to be clear when every observer is led to the same result when measuring the same physical entity. It should be stressed, however, that this is an empirical criterion perpetually subject to critique; thus physical entities which today are considered to be well defined may no longer be so in the future.

Hence, the physicist, from his observations of nature, obtains a set of numbers corresponding to the performance of some operations which are considered to be "objectively defined". Trying to organize such numbers coherently, the physicist often formulates "models".

In the attempt to organize coherently such numbers, the physicist formulates "models": i.e. he associates well-defined mathematical structures with his measurements, and he tries to establish a (small) number of mathematical relationship among them. From such relationships new ones logically follow, which reinterpreted through the model, used inversely, may serve to predict new relations between various empirical measurements.

The belief in the existence of good models motivated Galileo to write: "*Philosophy is written in the great book which is always open before our eyes (I mean the universe) but it cannot be understood unless one first learns the language and distinguishes the characters in which it is written. It is a mathematical language and the characters are triangles, circles and other geometrical figures, without which it cannot be understood by the human mind; without them one would vainly wonder through a dark labyrinth*".[1]

A mathematical model is considered satisfactory whenever it does not lead to contradictions with the experiments. If a contradiction occurs, the physicist dismisses the model as "wrong"; nevertheless, the mathematical construction built with it remains valid and is witness to an imperfect representation of nature.

Strictly speaking there is no model which is not wrong: only models that have not yet been shown to be wrong exist. However, all "serious" models (such as the dynamics of point masses, the theory of relativity, quantum mechanics, electromagnetism, thermodynamics, statistical mechanics, etc.) have led, and still lead, to the formulation of extremely interesting mathematical problems. Furthermore, it often happens that the analysis of the mathematical properties of a "wrong" model helps in the formulation of the new "more elaborate" model that the physicist tries to set up as a substitute.

A link between phenomena reality and mathematics can therefore be established as just described, through what has been called "a model". However,

[1] G.Galilei, *Il Saggiatore*, p. 232, [20].

it would be impossible to give a precise mathematical definition of the notion of a model because it is a rather empirical notion which can only be well understood through the analysis of several concrete cases.

1.2 An example of a Model

Consider the historically particularly important and significant case of the "mechanics of point masses". Its construction from empirical observations will be briefly and concretely analyzed, presenting it as a model of one or several point masses subject to forces.

The first statement (or "axiom", to use a mathematical term) says that the point masses are in a three-dimensional Euclidean space $\mathcal{R}^3$ in which any point can be represented by its three coordinates with respect to an orthogonal reference system $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$. The notation means that O is the origin and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the three orthogonal unit vectors pointing along the x, y, z coordinate axes, respectively.

Such an idealization has a clear mathematical meaning, but it appears to be unprovable in mathematical terms: it just renders the following empirical observation.

In practice, a point in space is determined by measuring (often only in principle and with the ruler method described in §1.1) its distance from three orthogonal walls. It is to be remarked that all such operations are ordinarily considered well defined.

A second statement (or "axiom") concerns "time" which, for the physicist, is the physical entity measured by a "clock" (classically described as a pendulum, although any more modern device will do as well). One assumes that time is an absolute "entity": in other words, one states that, at least in principle it is possible to associate with every point in space a clock mechanically identical at every point, and, furthermore, to coordinate ("synchronize") the clocks.

This means that if P, P' are two points and t, t' are two chosen time instants $t < t'$ it is then possible to send a signal from P towards P' leaving P at time t and reaching P' at time t' (as indicated by the local clocks in P and in P', respectively); while, vice versa, if $t > t'$, the above operation should be impossible.

A little thought makes it clear that the operational definition of a "system of synchronized clocks" is based on the empirical fact that it is possible to send signals with arbitrary speed. It is also clear that the notion of time is a phenomenological notion, far from being mathematically well posed.

Accepting the point of view so far discussed, one is led to say that the mathematical scheme, or model, representing the space-time continuum,where our observations take place, consists of a four-dimensional space: each of its points (x, y, z, t) represents a point seen in a Cartesian coordinate frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$

("laboratory") and observed at the instant t (as measured by the formerly introduced universal clocks).

Empirically, a point mass is any object which, at least as far as our observations are concerned, can be assimilated with a point in space (for instance, a planet or a star in the universe, a stone falling in a ravine, a ship sailing in the ocean, etc.). Such a point preserves its identity over the course of time; hence, it is possible to define its trajectory through a function of time $t \to \mathbf{x}(t)$, where $\mathbf{x}(t) = (x(t), y(t), z(t))$ is the vector whose components are the coordinates of the point at time t, in the chosen reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$.

Mathematically, a point mass moving in the reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ observed as t varies over an interval I is represented as a curve C in $\mathcal{R}^3$ by the vector equations $P(t) - O = \mathbf{x}(t),\ t \in I$; and the parameter t has the interpretation of time (i.e., it is called "time").

Given a point mass moving as t varies in I, one can associate with it its "velocity" at time $t \in I$. Operationally, velocity is defined by fixing $t_0 \in I$, finding the positions $P(t_0)$ and $P(t_0 + \varepsilon)$, and setting

$$\mathbf{v}(t_0) = \frac{P(t_0 + \varepsilon) - P(t_0)}{\varepsilon}, \tag{1.2.1}$$

where the parameter $\varepsilon > 0$ is to be chosen "suitably small" (according to well-defined criteria which, however, depend on the concrete cases). The mathematical model defines the point mass velocity at time $t_0 \in I$ as the derivative of the function $t \to \mathbf{x}(t)$ at $t = t_0$.

To complete the mathematical model of a point mass, it is important to define the "force" acting on it.

Operationally, the force acting at a given instant on the point mass consists of three scalar quantities which together define a vector $\mathbf{f}(t)$. The force acting on the point mass moving in $\mathcal{R}^3$ and observed in the frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ is measured through a "dynamometer" which is an instrument whose use is convenient to describe in a strongly idealized form. It is, basically, a suitably built spring which will be imagined as a very thin, light segment with a hook.

Consider a point mass moving in $\mathcal{R}^3$, with a velocity $\mathbf{v} = (v_x, v_y, v_z)$ relative to the reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ at time t_0. To measure the force acting upon it, hook it to the dynamometer to which the same velocity $\mathbf{v}$ has been imparted and which will be kept fixed during the measurement. Then try to adjust the spring length and direction so that the acceleration at time $t_0 + \varepsilon$ is 0, where $\varepsilon > 0$ is chosen "suitably small". (The empirical notion of acceleration and the corresponding mathematical model of it, as the second derivative with respect to t of the point position, is discussed along the same lines as the notion of velocity.)

The force is then the vector $\mathbf{f}$ whose direction is that of the dynamometer at time $t_0 + \varepsilon$, whose orientation is that parallel to hook but pointing away from it and whose modulus is the size of the spring elongation.

Summarizing: a point mass subject to forces and observed in a frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ in $\mathcal{R}^3$ as time varies within an interval I is, in its mathematical

model, described by a curve in seven-dimensional space: one of its points $(t, x, y, z, f_x, f_y, f_z)$ represents a point mass which at time t has coordinates (x, y, z) in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ and, in the same frame, is subject to a force (f_x, f_y, f_z). The curve representing this situation can be parameterized by the parameter t itself, as t varies in some time interval I; it shall also be assumed that in this parametric representation the functions $t \to (x(t), y(t), z(t))$ are twice continuously differentiable so that a mathematical definition of velocity and acceleration is meaningful.

1.3 The Laws of Mechanics

Once it is established what is meant by a point mass subject to forces and studied in a given frame of reference in $\mathcal{R}^3$ as the time varies in an interval I (briefly, "a point mass subject to forces"), it is possible to complete the mathematical model of the point mechanics. For this purpose, the "laws of dynamics" and their mathematical interpretation have to be discussed.

Experimentally, given a point mass, a simple relation is observed between its acceleration $\mathbf{a}$ at time t (in a given frame of reference) and the force $\mathbf{f}$ acting on it at that time (observed in the same frame). Such a relation is called the Second Law of Mechanics and establishes the existence of a constant $m > 0$, characteristic of the point mass and independent of the frame of reference used for the observations, such that:

$$m\mathbf{a} = \mathbf{f}. \tag{1.3.1}$$

This law introduces, via the properties of the differential equations, many relations among the quantities $\mathbf{x}, \mathbf{v}, t$, and such relations can sometimes be experimentally checked. For instance, if it is known a priori which force will act on the point mass whenever it is at the point (x, y, z) at time t with velocity (v_x, v_y, v_z), then, denoting such force as $\mathbf{f}(v_x, v_y, v_z, x, y, z, t) = \mathbf{f}(\mathbf{v}, \mathbf{x}, t)$, the differential equation

$$m\,\ddot{\mathbf{x}} = \mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, t) \tag{1.3.2}$$

allows the determination of the motion following an initial state, in which the velocity $\mathbf{v}_0$ and the position $\mathbf{x}_0$ are given at time t_0, at least for a small time interval around t_0 if $\mathbf{f}$ is a smooth function, see Chapter 2.

The First Principle of Mechanics postulates the existence of at least one reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, called "inertial frame", in $\mathcal{R}^3$ where a point mass "far" from the other objects in the universe appears to be subjected to a null force in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$. Such a frame is experimentally identified with a frame with origin in a fixed star and with axes oriented towards three more fixed stars. It is to such a frame that motion is often referred.

Of course the notions of "far" and of "fixed star" are empirical notions rather than mathematical ones.

Mathematically, the first principle is used to grant to a particular frame of reference in the space-time continuum a privileged role and to define the "absolute force" or the "true force" as that acting on the point mass in this frame. This frame has to be chosen once and for all and is called the "fixed reference frame" (as opposed to "moving reference frame").

It is possible and sometimes convenient to introduce frames whose origin and axes vary with time with respect to the "fixed" frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$: $(0(t); \mathbf{i}(r), \mathbf{j}(t), \mathbf{k}(t))$.

Since $\mathbf{f} = m\mathbf{a}$, it follows that if the moving frame is in uniform rectilinear translational motion with respect to the fixed frame, then the force acting upon the point is the same whether observed in the fixed frame or in the moving frame: hence, in this moving frame, the "inertia principle", i.e., the first principle, is valid: a point mass which is "very far" from the other objects in the universe is subject to a null force, since the acceleration is the same in the two frames. All frames in rectilinear uniform motion with respect to a fixed frame are called "inertial frames".

The mathematical model of a point mass with mass m subject to forces and obeying the laws of dynamics is then, simply, a point mass subject to forces, in the sense of the preceding section, and such that the relation

$$m\,\mathbf{a} = \mathbf{f} \tag{1.3.3}$$

holds and, furthermore, $\mathbf{f}$ is a function of the point velocity, position, and time; i.e., the following relation holds:

$$\mathbf{f} = \mathbf{f}(\mathbf{v}, \mathbf{x}, t). \tag{1.3.4}$$

Clearly, from such a mathematical viewpoint (where $\mathbf{f}$ is imagined as given a priori), the first principle is deprived of its deep physical meaning.

An important extension of the point mass model is a model for the mechanics of a "system of N point masses". Mathematically, such a system consists of N point masses with mass $m_1, \ldots, m_N$, in the above sense, satisfying the Third Principle of Mechanics. This means that it should be possible to represent the force $\mathbf{f}_i$ acting on the i-th point as

$$\mathbf{f}_i = \sum_{j \neq i} \mathbf{f}_{j \to i}, \tag{1.3.5}$$

where $\mathbf{f}_{j \to i}$ are such that
(a) $\mathbf{f}_{j \to i} = -\mathbf{f}_{i \to j}$, $j, i = 1, 2, ..., N$, $i \neq j$;
(b) $\mathbf{f}_{j \to i}$ is parallel to $P_j - P_i$, i.e., to the line joining the positions P_i and P_j of the i-th and j-th points;
(c) $\mathbf{f}_{j \to i}$ depends solely upon the positions and velocities of the it-h and j-th points and on time:

$$\mathbf{f}_{j \to i} \equiv \mathbf{f}_{j \to i}(\mathbf{v}_j, \mathbf{v}_i, P_j, P_i, t). \tag{1.3.6}$$

This assumption corresponds to a precise empirical fact: it is possible to define operationally what should be understood by $\mathbf{f}_{j\to i}$ "the force exerted by the point P_j on the point P_i".

For instance, the force $\mathbf{f}_{j\to i}$ could be measured as follows: one measures, in the given inertial frame of reference, the force $\mathbf{f}_i$, acting on i and then one measures, after removing the point j from the system, the new force acting on the i-th point, obtaining the result $\mathbf{f}_i^{(j)}$; then one sets

$$\mathbf{f}_{j\to i} = \mathbf{f}_i - \mathbf{f}_i^{(j)}. \tag{1.3.7}$$

The Third Principle of Mechanics arises from the experimental observation that $\mathbf{f}_{j\to i} = -\mathbf{f}_{i\to j}$, that $\mathbf{f}_{j\to i}$ is parallel to $P_j - P_i$, that the total force acting on a singe point mass is the sum of the forces exerted on it by the other system points (in the sense of vectors addition) if observed in an inertial frame of reference, and, finally, that $\mathbf{f}_{j\to i}$ depends only upon the positions and velocities of the points involved and, possibly, on time.

Physics often places still more requirements and restrictions upon the laws of force which can be used to give a more detailed specification of a mechanical system model. However, they do not have a general character comparable to the three principles but, rather, are statements explaining which laws of force are to be considered a good model under given circumstances. For instance, two point masses "without structure" (this is, again, an empirical notion which we refrain from elucidating) attract each other with a force of intensity mm'/kr^2, where r is the distance between the points, m and m' are their masses, and k is a universal constant. If the structure of the two points can be summarized by saying that they have an "electric charge e" (a new empirical notion), the mutual force will be the vector sum of the above-described gravitational force and of a repulsive force with intensity $k'e^2/r^2$, where k' is another universal constant.

The principles of mechanics already place enough restrictions upon the nature of the forces admissible in mechanical problems: therefore it is convenient and interesting to examine their implications before passing to the analysis of special models obtained by concretely specifying the "force laws", i.e., the functions giving the forces in terms of the points positions and velocities and of time.

It should be stressed, and this is a general comment on the mathematical models for physical phenomena, that the mathematical model is always "poorer" than the physical reality that it tries to imitate. For instance in the above mathematical model for mechanics, the first principle loses its meaning. Another example, implicit in the above discussion, is the following.

To give an operational meaning to the notions of position, speed, force, etc., it must be possible to repeat "identical" experiments several times (e.g., see the position measurement in §1.1 by repeating the measurement operations. However, time inexorably flows away, and this is impossible. Physically, this difficulty is avoided by the "principle of homogeneity of space-time" which

says that experiments starting at any time in any space location will yield the same results if the points involved are in the same relative positions and situations.

In the mathematical model for mechanics just described, the necessity of understanding the above problems does not arise, nor do many other similar problems which the reader will easily think of.

Usually it is possible to complicate the models in order to imbue them with any given number of physical facts: but an analysis of this type of questions would lead us beyond the scope of this book.

In any case, a decision is always needed on where to put a stop to the process of model improvement, which would otherwise hopelessly continue ad infinitum. We must recall that we have the more down-to-earth, and more interesting, problem of obtaining some concrete prediction algorithms for our observations of nature.

1.4 General Thoughts on Models

In this book more abstract schematization processes concerning empirically observed phenomena will be met (e.g., when we discuss the notion of an "observable" or of a "vibrating string"). In such cases, however, the details of the construction of the mathematical model will not be repeated: a very common practice based on the idea that the very words used to designate well-defined mathematical objects will implicitly define the model.

It is such a practice, or better, its imperfect understanding, which sometimes causes misunderstandings between physicists and mathematicians and provokes allegations of non-rigorous use of mathematics.

It is important to realize that when the physicist speaks in mathematical terms he is by no means attributing to them the same rigid meaning that a mathematician would assume for them. Rather he is using this language to help himself in the formulation of a model which, once well defined, he shall rigorously treat (since he believes, or at least hopes, that the book of nature is written in mathematical characters).

Possibly logically non rigorous steps or apparently wild mathematical approximations in a physicist's argument should always be interpreted as further complications or, better, refinements of the model that the physicist is trying to build.

In the hectic development of research, a physicist often modifies a model while using it, or he modifies the mathematical meaning of the objects and entities which belong to the model without changing their names (otherwise, a dictionary would not suffice). He does this because his main interest is in the construction of models and only secondarily in its mathematical theory, often considered trivial for his needs.

To avoid excessively pedantic discussions, we shall adhere, in the following, to the well-established practice of avoiding the physical analysis necessary to

the construction of a model and shall leave it to the reader to imagine such an analysis via the suggestive names used for the various mathematical entities (with the exception of a few important cases). In any case, this book is devoted to the mathematical, rather than physical aspects, of mechanical problems.

Bibliographical Comment. It is very useful to study at least the definition and the laws of motion in the *Philosophiae Naturalis Principia Mathematica* by I. Newton, [37], to understand exactly the Newtonian formulation of mechanics and its modernity. To avoid "reading too much", i.e., to avoid interpreting these immortal pages in too modern a way, it is a good idea to read the paper *Essays on the history of mechanics* by C. Truesdell, pp. 85-137 ([48]). The reading of the first two chapters of the work by E. Mach, [31],) will be a very useful and stimulating complement to the first three chapters of this book.

2

Qualitative Aspects of One-Dimensional Motion

2.1 Energy Conservation

Consider a point mass, with mass m, on the line $\mathcal{R}$ and subject to a force law depending uniquely on its position. Therefore, a force law $\xi \to f(\xi)$ is, given $\xi \in \mathcal{R}$, which we shall suppose to be of class C^∞, associating with every point ξ on the line $\mathcal{R}$ the component $f(\xi)$ of the force acting on the point when it happens to occupy the position ξ.

A "motion" of the point mass, observed as t varies in an interval I, is a function $t \to x(t)$, $t \in I$, of class $C^\infty(I)$ such that

$$m\,\ddot{x}(t) = f(x(t)), \qquad \forall\, t \in I \tag{2.1.1}$$

The "energy conservation theorem" follows by multiplying Eq. (2.1.1), side by side, by $\dot{x}(t)$:

$$m\,\dot{x}\,\ddot{x} = \dot{x}\,f(x), \tag{2.1.2}$$

omitting, as will often be done, the explicit mention of the t-dependence. Then, defining the functions,

$$\eta \to T(\eta) \stackrel{def}{=} \frac{1}{2} m\,\eta^2, \qquad \xi \to V(\xi) \stackrel{def}{=} -\int^{\xi} f(\xi')\,d\xi', \tag{2.1.3}$$

it is

$$\frac{d}{dt}T(\dot{x}) = m\,\dot{x}\,\ddot{x}, \qquad \frac{d}{dt}V(x) = -f(x)\,\dot{x} \tag{2.1.4}$$

so that Eq. (2.1.2) becomes

$$\frac{d}{dt}(T(\dot{x})+V(x))=0 \tag{2.1.5}$$

This implies a constant E can be associated with every motion $t \to x(t)$, $t \in I$, depending on the motion under consideration and such that

$$T(\dot{x}(t))+V(x(t))=E, \qquad \forall\, t \in I. \tag{2.1.6}$$

The expressions $T(\dot{x})$ and $V(x)$ are respectively called the "kinetic energy" and the "potential energy" and Eq. (2.1.6) has to be read as follows: "in every motion developing under the action of a force with potential energy V, the sum of the kinetic energy and potential energy is a constant". This constant is given the name "total energy" of the considered motion. The "qualitative theory" of Eq. (2.1.1) is concerned with the analysis of the properties of the motion verifying Eq. (2.1.1), which are valid independently of the choice of f, at least for vast classes of functions f. The energy conservation is a first example of a qualitative property.

Observations. The energy conservation goes back at least to Huygens; afterwards, it was used by J. and D. Bernoulli together with the law of conservation of linear momentum (Descartes) (see [48], p. 105 and following).

Eq. (2.1.6) implies an expression for the velocity:

$$\dot{x}(t)=\pm\big(\frac{2}{m}(E-V(x(t)))\big)^{\frac{1}{2}}, \qquad t \in I \tag{2.1.7}$$

This relation, which will be used and discussed in §2.6, allows the reduction of the determination of the evolution law $t \to x(t)$, $t \in I$, "time law", to an area-computation problem for a planar figure, "quadrature". In fact, supposing $\dot{x} > 0$, it yields:

$$t=\int_{x(0)}^{x(t)} \frac{d\xi}{\sqrt{\frac{2}{m}(E-V(\xi))}}\, d\xi \tag{2.1.8}$$

when $I \supset [0,t]$.

Hence, the area under the graph of the curve with equation $\xi \to T(\xi) = (\frac{2}{m}(E-V(\xi)))^{-\frac{1}{2}}$ above the interval $[x(0), x(t)]$ is the time that the point needs to reach $x(t)$, starting from $x(0)$ at time 0 with positive speed and energy E, at least for small t (i.e., as long as $\dot{x} > 0$).

Newton "reduced to quadratures" the simplest problems of motion without explicitly using energy conservation ([37], for instance Book I, Propositions XXXIX, XLI, LIII, LVI, etc.).

2.2 General Properties of Motion. Uniqueness

In the preceding §2.1, a motion developing, under the action of a force f, in a time interval I was supposed to be given. We can ask which further properties of a particular motion allow us to select it from among all motions which, in the same time interval I, take place under the action of the same force.

One can even preliminarily ask whether, given an interval I, there exist any motions, i.e., C^∞ solutions of Eq. (2.2.1) thought of as an equation for $t \to x(t)$, $t \in I$.

In view of the importance of such questions, before proceeding in the analysis of Eq. (2.1.1), some attention will be devoted to the general problem of the existence, uniqueness, and regularity of the solutions of differential equations in $\mathcal{R}^d$.

Eq. (2.1.1), thought of as a "second-order" differential equation in $\mathcal{R}^1$, is equivalent to a "first-order" equation in $\mathcal{R}^2$: it suffices to write it as

$$\dot{x}(t) = y(t), \qquad \dot{y}(t) = f(x(t)), \tag{2.2.1}$$

where Eq. (2.2.1) is an equation for the unknown C^∞ function $t \to (x(t), y(t))$ defined on I and with values in $\mathcal{R}^2$.

More generally, consider an arbitrary "s-th order" differential equation in $\mathcal{R}^d$, $s = 0, 1, \ldots$, like

$$\frac{d^s\mathbf{x}(t)}{dt^s} = \mathbf{f}(\frac{d^{s-1}\mathbf{x}(t)}{dt^{s-1}}, \ldots, \frac{d\mathbf{x}(t)}{dt}, \mathbf{x}(t), t), \tag{2.2.2}$$

with $t \in I$, where f is an $\mathcal{R}^d$-valued C^∞ function defined on $\mathcal{R}^d \times \mathcal{R}^d \times \mathcal{R}$ and $t \to \mathbf{x}(t)$ is an unknown $\mathcal{R}^d$-valued C^∞ function on I. The latter equation may be thought of as a first-order equation in $\mathcal{R}^d$ by setting

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \ \mathbf{y}_1, & \frac{d\mathbf{y}(t)_1}{dt}(t) &= \mathbf{y}_2, \ldots \\ \frac{d\mathbf{y}(t)_{s-2}}{dt} &= \mathbf{y}_{s-1}, & \frac{d\mathbf{y}(t)_{s-1}}{dt} &= \mathbf{f}(\mathbf{y}_{s-1}(t), \ldots \mathbf{y}_1(t), \mathbf{x}(t), t) \end{aligned} \tag{2.2.3}$$

and then considering Eq. (2.2.3) as an equation for the C^∞ function $t \to (\mathbf{x}(t), \mathbf{y}_1(t), \ldots .\mathbf{y}_{s-1}(t))$ defined on the interval I and with values in $\mathcal{R}^d \times \ldots \times \mathcal{R}^d = \mathcal{R}^{d\,s}$.

Eq. (2.2.2) is the most general differential equation that will be met in this book. By virtue of the preceding remark, it will then suffice, for our purposes, to study first-order differential equations in $\mathcal{R}^d$ having the form

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \qquad t \in I, \tag{2.2.4}$$

It will be convenient to introduce a precise convention about what a differential equation is or about what one of its solutions is.

1 Definition. *Given an $\mathcal{R}^d$-valued function $\mathbf{F} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$, the expression (2.2.4), denoted, for short, $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$, will be called a "differential equation on $\mathcal{R}^d$ associated with $\mathbf{F}$".*
A "$C^{(k)}$ solution", $k > 1$, of Eq. (2.2.4) on the interval I, closed or open or semi open, will be a $C^{(k)}$ function which turns Eq. (2.2.4) into an identity when substituted into it.[1] *A "solution" of Eq. (2.2.4) for $t \in I$ is a C^∞ solution. The solutions of Eq. (2.2.4) will often be called "motions".*

Let us first examine the uniqueness problem for the solutions of Eq. (2.2.4).

1 Proposition. *Let $(\boldsymbol{\xi}, t) \to \mathbf{F}(\boldsymbol{\xi}, t)$ be an $\mathcal{R}^d$-valued C^∞ function on $\mathcal{R}^d x\mathcal{R}$. Given $a > 0, b > 0, t_0 \in \mathcal{R}$, let $t \to \mathbf{x}(t)$ be a $C^{(1)}$ solution of Eq. (2.2.4) on $J = [t_0 - a, t_0 + b]$:*
(i) the function $t \to \mathbf{x}(t)$ is in $C^\infty(J)$;
(ii) if $t \to \mathbf{y}(t)$ is another solution of Eq. (2.2.4) on J and if $\mathbf{y}(t_0) = \mathbf{x}(t_0)$, then $\mathbf{x}(t) = \mathbf{y}(t), \forall\, t \in J$.

Observations.

(1) This proposition applied to Eq. (2.2.2) via Eq. (2.2.3) tells us that two $C^{(s)}$ solutions of an s-th order differential equation in $\mathcal{R}^d$ for $t \in J$ coincide if and only if at time $t_0 \in J$ ("initial time") they have the same first $(s-1)$ derivatives ("equal initial data"). When Eq. (2.2.2) is the equation governing a physical motion in $\mathcal{R}^d$, it is $s = 2$; this means that the motion is uniquely determined, if existing at all, by its initial position $\mathbf{x}(t_0)$ and by its initial velocity $\dot{\mathbf{x}}(t_0)$, i.e., as one says, by its initial "act of motion" $\dot{\mathbf{x}}(t_0)$.
(2) It would appear that it might be interesting or important to know if, by specifying properties of the solutions of Eq. (2.2.2) other than the just-mentioned initial data at some initial time, the solution verifying such properties is uniquely determined [2], if existing at all. The uniqueness criterion that we chose above for illustration purposes, Proposition 1, has been selected only because it quickly leads to a simple answer and because it is one of the uniqueness criteria which are most useful in many applications.
(3) From the proof it will appear that if $\mathbf{F}$ had been only supposed to be of class $C^(k)$, $k \geq 1$, then uniqueness would have followed in an equal way. The regularity of $t \to \mathbf{x}(t), t \in J$, could also be deduced in this case, but one would only obtain that $t \to \mathbf{x}(t)$ is a $C^{(k+1)}$ function.

Proof. By integrating both sides of Eq. (2.2.4) and by setting $\mathbf{x}_0 = \mathbf{x}(t_0) = \mathbf{y}(t_0)$, we get:

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{F}(\mathbf{x}(\tau), \tau)\, d\tau, \qquad t \in J, \tag{2.2.5}$$

[1] We shall see that every $C^{(k)}$ solution, $k > 1$, is automatically a C^∞ solution, if $\mathbf{F} \in C^\infty$.
[2] For instance, we can ask the following question. Consider Eq. (2.2.2) with $s = 2$ and lei t_1, t_2 be two times and $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{R}^d$ be two positions. Is the motion [solution of Eq. (2.2.2)] leading from $\mathbf{x}_1$ to $\mathbf{x}_2$ as time elapses from t_1 to t_2 (assuming that one such motion, at least, exists) unique? We shall see that the answer to this question will, in general, be no.

and, similarly, since also $t \to \mathbf{y}(t)$ is a solution of Eq. (2.2.4):

$$\mathbf{y}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{F}(\mathbf{y}(\tau), \tau)\, d\tau, \qquad t \in J. \tag{2.2.6}$$

Hence,

$$\mathbf{x}(t) - \mathbf{y}(t) = \int_{t_0}^{t} (\mathbf{F}(\mathbf{x}(\tau), \tau) - \mathbf{F}(\mathbf{y}(\tau), \tau))\, d\tau. \tag{2.2.7}$$

To prove (ii) the procedure that will be followed is very interesting since it obviously goes beyond the particular result that we wish to obtain.

Informally, the argument is the following: the difference $|\mathbf{x}(t) - \mathbf{y}(t)|$ is, by Eq. (2.2.7), about $|t - t_0|\, |\mathbf{F}(\mathbf{x}(t), t) - \mathbf{F}(\mathbf{y}(t), t)|$, if $t \sim t_0$; however, the increment $|\mathbf{F}(\mathbf{x}(t), t) - \mathbf{F}(\mathbf{y}(t), t)|$ is proportional, by Lagrange's theorem, to the increment of the argument of $\mathbf{F}$, i.e., to $C\, |\mathbf{x}(t) - \mathbf{y}(t)|$, where C is an estimate of the first derivatives of $\mathbf{F}$. Hence, Eq. (2.2.7) implies that $|\mathbf{x}(t) - \mathbf{y}(t)|$ and $C|t - t_0|\, |\mathbf{x}(t) - \mathbf{y}(t)|$ are about equal if $t \sim t_0$, and this, in turn, implies that $|\mathbf{x}(t) - \mathbf{y}(t)| = 0$ for t dose to t_0 because for $t \sim t_0$, one has $C|t - t_0| < 1$.

To estimate the integrand of Eq. (2.2.7) let $S \subset \mathcal{R}^d$ be a sphere with so large a radius that it contains all the values $\mathbf{x}(\tau), \mathbf{y}(\tau), \forall\, \tau \in J$, and let

$$M_S = \max_{\boldsymbol{\xi} in S, \tau \in J} \sum_{i,j=1}^{d} \Big| \frac{\partial F^{(i)}}{\partial \xi_j} \Big| \tag{2.2.8}$$

where $F^{(i)}(\boldsymbol{\xi}, t)$ is the i-th component of the vector $\mathbf{F}(\boldsymbol{\xi}, t) = (F^{(1)}(\boldsymbol{\xi}, t), \ldots, F^{(d)}(\boldsymbol{\xi}, t)) \in \mathcal{R}^d$. Then, from Taylor's formula:

$$|\mathbf{F}(\mathbf{x}(\tau), \tau) - \mathbf{F}(\mathbf{y}(\tau), \tau)| \le M_S\, |\mathbf{x}(\tau) - \mathbf{y}(\tau)|. \tag{2.2.9}$$

Inserting this inequality into Eq. (2.2.7), yields

$$|\mathbf{x}(t) - \mathbf{y}(t)| \le M_S \int_{t_0}^{t} |\mathbf{x}(\tau) - \mathbf{y}(\tau)|\, d\tau \tag{2.2.10}$$

Let $M(t) = \max_{t_0 \le \tau \le t} |\mathbf{x}(\tau) - \mathbf{y}(\tau)|$, $t \in [t_0.t_0 + b]$; then Eq. (2.2.10) implies $|\mathbf{x}(t) - \mathbf{y}(t)| \le M_S\, M(t)\, |t - t_0|, \forall\, t \in [t_0, t_0 + b]$.

Since $M(t)$ is monotonic nondecreasing and since this inequality holds for all $t \in [t_0, t_0 + b]$, one easily finds that

$$M(t) \le M_S |t - t_0|\, M(t), \qquad \forall\, t \in [t_0, t_0 + b] \tag{2.2.11}$$

which implies $M(t) = 0$ for $|t - t_0| < M_S^{-1}, t \in [t_0, t_0 + b]$.

Hence, $\mathbf{x}(t_0 + M_S^{-1}) = \mathbf{y}(t_0 + M_S^{-1})$, if $t_0 + M_S^{-1} < t_0 + b$, and the argument can be repeated, replacing t_0 by $t_0 + M_S^{-1}$, to show that $M(t) = 0$ for $t \in [t_0, t_0 + 2M_S^{-1}]$ if $t_0 + 2M_S^{-1} < t_0 + b$, etc., so that $M(t) = 0$ for $t \in [t_0, t_0 + b]$. For $t \in [t_0 - a, t_0]$, one proceeds likewise.[3]

---

[3] Alternatively, Eq. (2.2.10) could be iterated n times to yield, if $\mu = \max |\mathbf{x}(\tau) - \mathbf{y}(\tau)|$, $\tau \in [t_0 - a, t_0 + b]$:

To check (i), i.e., that $t \to \mathbf{x}(t)$ is a C^∞ function on J, remark that if $t \to \mathbf{x}(t)$ is a $C^{(1)}(J)$ function, then Eq. (2.2.4) implies that $t \to \dot{\mathbf{x}}(t)$ is in $C^{(1)}(J)$, being a composition of a C^∞ function with a $C^{(1)}$ function; furthermore, by differentiating Eq. (2.2.4):

$$\ddot{\mathbf{x}}(t) = \sum_{i=1}^{d} \frac{\partial \mathbf{F}}{\partial \xi_i}(\mathbf{x}(t), t) \cdot \dot{\mathbf{x}}^{(i)} + \frac{\partial \mathbf{F}}{\partial t}(\mathbf{x}(t), t) \tag{2.2.12}$$

which, in turn, implies that $t \to \ddot{\mathbf{x}}(t)$ is a $C^{(1)}$ function by the same argument as above. Then, by differentiating Eq. (2.2.12), one finds that $\dddot{\mathbf{x}}(t)$ is a $C^{(1)}$ function on J, etc. mbe

2.2.1 Problems for §2.2

1. If $t \to \mathbf{x}(t), t \geq 0$, solves $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and $\mathbf{x}(0) = \mathbf{x}(T)$ for some $T > 0$, then $\mathbf{x}(t) = \mathbf{x}(t+T), \forall\, t > 0$; assume $\mathbf{f} \in C^\infty(\mathcal{R}^d)$. Would this also be true if $\mathbf{f} \in C^1(\mathcal{R}^d)$? (Hint: Use uniqueness).

2. The property of the preceding problem is not valid when the differential equation right-hand side is explicitly time dependent (i.e., $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, and $\partial \mathbf{f}/\partial t \neq \mathbf{0}$, the "non autonomous case"). Find an example.

3. Let $\mathbf{f}(\mathbf{x}, t)$ be such that $\mathbf{f}(\boldsymbol{\xi}, t) = \mathbf{f}(\boldsymbol{\xi}, t+T)$ for some $T > 0$ and for all $\boldsymbol{\xi} \in \mathcal{R}^d$. Suppose that $t \to \mathbf{x}(t)$ is a solution of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ such that for some integer $m > 0$, one has $\mathbf{x}(0) = \mathbf{x}(mT)$, then $\mathbf{x}(t) \equiv \mathbf{x}(t+mT), \forall\, t \geq 0$. (Hint: Use uniqueness.)

4. Consider the equation $\dot{x}(t) = \ell(t)\, x(t)$ with $\ell \in C^\infty(\mathcal{R})$. Show that if $t \to x(t)$ and $t \to y(t)$ are two solutions for $t \in J$ and if $x(t) \not\equiv 0$, there exists a constant A such that $y(t) \equiv Ax(t), \forall\, t \in J$.

5. If the function ℓ of the Problem 4 is periodic with period $T > 0$ and $t \to x(t) \not\equiv 0$, is one of its solutions then also $t \to x(t+T)$ is a solution. Hence, $\exists \lambda \neq 0$ such that $x(t+T) = \lambda x(t)$. Show that $\lambda > 0$. (Hint: Otherwise either $\lambda = 0$ and $x(T) = 0$, hence $x(t) \equiv 0$ (by uniqueness on $[0, +\infty)$), or $\lambda < 0$ and there would be $\overline{t} \in (0.T]$ where $x(\overline{t}) = 0$: hence, again, $x(t) = 0$ by uniqueness.)

6. The most general solution $t \to y(t)$, $t \in \mathcal{R}_+$, of the equation in Problem 4, with ℓ periodic with period T has the form $y(t) = A\lambda^{t/T} z(t)$, where $z \in C^\infty(\mathcal{R}_+)$ is T-periodic.

7.* Consider the equation $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$ in $\mathcal{R}^d$, where $t \to \mathbf{L}(t)$, $t \in \mathcal{R}$, is a $d \times d$-matrix valued C^∞ function. Consider d solutions $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}$ for $t \in I = [a, b]$ and call them "independent" if $\exists\, t_0 \in I$ such that the d vectors $\mathbf{x}^{(1)}(t_0), \ldots, \mathbf{x}^{(d)}(t_0)$ are linearly independent. Show that, if $t \in I$, then also $\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(d)}(t)$ are linearly independent whenever they are such for $t = t_0$ and, furthermore, any solution $t \to \mathbf{y}(t)$, $t \in I$, can be represented as $\mathbf{y}(t) = \sum_{j=1}^{d} A_j \mathbf{x}^{(j)}(t)$, $\forall\, t \in I$. (Hint: If for $t = \overline{t}$, the d vectors were not independent,

---

$$\begin{aligned} |\mathbf{x}(t) - \mathbf{y}(t)| &< M_S^n \int_{[t_0,t]} d\tau_1 \int_{[t_0,\tau_1]} d\tau_3 \ldots \int_{[t_0,\tau_{\nu-1}]} d\tau_n\, |\mathbf{x}(\tau_n) - \mathbf{y}(\tau_n)| \\ &\leq M_S^n \mu \int_{[t_0,t]} d\tau_1 \int_{[t_0,\tau_1]} d\tau_3 \ldots \int_{[t_0,\tau_{\nu-1}]} d\tau_n \; = M_S^n \mu \frac{|t - t_0|^n}{n!} \leq M_S^n \mu \frac{(a+b)^n}{n!} \end{aligned}$$

so that $\mathbf{x}(t) - \mathbf{y}(t) \equiv 0$ since n is arbitrary and it can be let to $+\infty$.

one could find constants $\overline{A}, \ldots, \overline{A}_d$, not all equal to zero, such that $\sum_{j=1}^d \overline{A}_j \mathbf{x}^{(j)}(\overline{t}) = \mathbf{0}$; hence, by linearity and uniqueness, $\sum_{j=1}^d \overline{A}_j \mathbf{x}^{(j)}(t) = 0, \forall t \in I$ which contradicts the independence for $t = t_0$.)

8. Show that Problem 7 implies that, given d solutions $t \to \mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(d)}(t)$, $t \in I$, to $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$, the matrix $W(t)$ ("Wronskian matrix" of $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}$) defined by

$$W_{ij}(t) = x_j^{(i)}(t), \qquad i, j = 1, 2, \ldots, d, \;\; t \in I$$

has a determinant $w(t)$ non vanishing for $t \in I$ if and only if $\exists t_0 \in I$ such that $w(t_0) \neq 0$. (*Hint.* By linear algebra, this is just another way of phrasing Problem 7: d vectors are linearly independent if and only if the "determinant of their components" is not zero.)

9. Using the determinant differentiation rule, by rows, show that

$$\frac{d}{dt} w(t) \equiv \frac{d}{dt} \det W(t) = \Big(\sum_{i=1}^d \ell_{ij}(t) \Big)\, w(t);$$

hence, if $\sum_{i=1}^d \ell_{ij}(t) = \ell(t)$, one has $w(t) = w(t_0)\, e^{\int_{t_0}^t \ell(\tau) d\tau}$.

10. In the context of Problem 8, suppose that the matrix function $t \to \mathbf{L}(t)$, $t \in \mathcal{R}$, is periodic with period $T > 0$, i.e., $t \to \ell_{ij}(t)$, $i, j = 1, \ldots, d$ are $T-$periodic functions. Let $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}$ be d linearly independent solutions for $t > 0$. Then there exist d^2 constants $A_j^{(i)}$, $i, j = 1, \ldots, d$, such that

$$\mathbf{x}^{(i)}(t+T) = \sum_{j=1}^d A_j^{(i)} \mathbf{x}^{(j)}(t), \qquad t \geq 0.$$

Show that $\det W(T)/\det W(0) = w(T)/w(0) = \det A \neq 0$.

11. Suppose that the matrix A is similar, via a real nonsingular matrix S, to a real diagonal matrix Λ, $\Lambda_{ij} = \lambda_i\, \delta_{ij}$, $i, j = 1, \ldots, d$: $SAS^{-1} = \Lambda$. In the context of Problem 10, define

$$\mathbf{y}^{(i)}(t) = \sum_{j=1}^d S_{ij} \mathbf{x}^{(j)}(t).$$

Show that $\mathbf{y}^{(1)}, \ldots . \mathbf{y}^{(d)}$ are linearly independent solutions, $\lambda_1, \ldots, \lambda_d \neq 0$, and

$$\mathbf{y}^{(i)}(t+T) = \lambda_i\, \mathbf{y}^{(i)}(t), \qquad t \geq 0$$

12. Suppose that A is a matrix similar to a diagonal matrix Λ via a complex nonsingular matrix S. Show that $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(d)}$, defined as in the preceding problem, are complex solutions of $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$ and that $\mathbf{y}^{(i)}(t+T) = \lambda_i\, \mathbf{y}^{(i)}(t)$, $\forall t \geq 0$. (For applications, recall that from linear algebra (see Appendix E), a sufficient condition for the similarity between A and a diagonal matrix $\Lambda_{ij} = \lambda_i\, \delta_{ij}$ is that the roots $\lambda_1, \ldots, \lambda_d$ of the secular equation $\det(A - \lambda) = 0$ are pairwise different.)

13. Given the assumptions of Problems 10,11 and supposing $\lambda_1, \ldots, \lambda_d > 0$, show that the most general solution to $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$ has the form

$$\mathbf{x}(t) = \sum_{j=1}^d \alpha_j \lambda_j^{t/T} \mathbf{z}^{(j)}(t)$$

where the functions $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(d)}$ are d C^∞ functions periodic with period T, and $\alpha_1, \ldots, \alpha_d$ are arbitrary constants. (Hint: Let $\mathbf{z}^{(i)}(t) = \lambda_i^{-t/T} \mathbf{y}^{(i)}(t)$.)

14. Suppose that for every nonzero complex number λ, there exists a C^∞ function $t \to \gamma(t)$, $t \in \mathcal{R}$, such that $\gamma(t+t') = \gamma(t)\gamma(t')$, $\gamma(0) = 1$, $\gamma(T) = \lambda^{-1}$, $\gamma(t) \neq 0 \,\forall\, t \in \mathcal{R}$; then the conclusions of Problem 13 would hold, replacing $\lambda^{-t/T}$ by $\gamma(t)$, without the assumption $\lambda_j > 0$, $j = 11, \ldots, d$, under the only assumption $\det A \neq 0$. See also the following problem.

15. Let $\lambda \in \mathcal{C}$, $\lambda^{-1} = \varrho\,(\cos\theta + i\,\sin\theta, \varrho > 0$, $\theta \in [0, 2\pi]$. Define $\gamma(t) = \varrho^{t/T}(\cos\frac{t}{T}\theta + i\,\sin\frac{t}{T}\theta)$. Show that $\gamma(0) = 1, \gamma(t)\gamma(t+t') = \gamma(t+t')$, $\gamma(T) = \lambda^{-1}, \gamma(t) \neq 0$, $\forall\, t \in \mathcal{R}$ (e.g., $(-1)^{t/T} = \cos\frac{t}{T}\pi + i\,\sin\frac{t}{T}\pi$).

Observations to Problems 8-15.

We shall see that there always exist d linearly independent solutions to $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$. However, the existence of S is a restrictive condition. When such an S does not exist, it is possible to show that the most general solution to $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$, with $\mathbf{L}$ periodic with period $T > 0$ and C^∞, can be written in the form

$$\mathbf{x}(t) = \sum_{j=1}^{p} \sum_{k=0}^{\delta(j)-1} \alpha_{jk} \lambda_j^{t/T} t^k z^{(j)}(t),$$

where $\sum_{j=1}^{p} \delta(j) = d$, and $\delta(j)$, λ_j are suitably chosen, and $t \to \mathbf{z}^{(j)}(t), t \geq 0$, are C^∞ functions periodic with period T and possibly complex valued (when λ_j are not positive and $\lambda_j^{t/T}$ is interpreted as explained in Problem 15), and α_{jk} are arbitrary constants (see [38], for instance, Vol. 1, pp. 63-68,).

16. Consider a differential equation $\ddot{x} + a(t)\dot{x} + b(t)x = 0, t \in \mathcal{R}$, $a, b \in C^\infty(\mathcal{R})$. After reducing it to a first-order system of two differential equations in $\mathcal{R}^2$, interpret the results of Problems 7-15 in terms of its solutions. Show first that the matrix $W(t)$ associated with this system is expressed in terms of two of its solutions $t \to x^{(1)}(t)$ and $t \to x^{(2)}(t)$ as

$$W(t) = \begin{pmatrix} x^{(1)}(t) & \dot{x}^{(1)}(t) \\ x^{(2)}(t) & \dot{x}^{(2)}(t) \end{pmatrix} \text{ and } \dot{w}(t) = a(t)w(t).$$

17.* Extend Problem 16 to the case of the sth-order differential equation in $\mathcal{R}$:

$$\frac{d^s x}{dt^s} + \sum_{j=0}^{s-1} a_j(t) \frac{d^j x}{dt^j}, \qquad t \in \mathcal{R}.$$

2.3 General Properties of Motion. Existence

An existence problem for the solutions of Eq. (2.2.4), hence of Eq. (2.2.2), naturally associated with the uniqueness property given in Proposition 1, §2.2, is solved by the following proposition:

2 Proposition. *Let* $\mathbf{F}$ *be an* $\mathcal{R}^d$*-valued function in* $C^\infty(\mathcal{R}^d \times \mathcal{R})$. *Let* $\mathbf{x}_0 \in \mathcal{R}^d$ *and* $t_0 \in \mathcal{R}$. *Let* $S(\boldsymbol{\xi}_0, \varrho)$ *be the closed ball in* $\mathcal{R}^d$ *with center* $\boldsymbol{\xi}_0$ *and radius* ϱ. *Let* $\theta > 0$. *There exists* $T_{\varrho,\theta} > 0$ *and a solution of Eq. (2.2.4), i.e.,* $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$, *defined for* $t \in [t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$ *and of class* C^∞ *such that:*

$$\mathbf{x}(t_0) = \boldsymbol{\xi}_0, \quad \mathbf{x}(t) \in S(\boldsymbol{\xi}_0, \varrho), \quad t \in [t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]. \tag{2.3.1}$$

Furthermore, if one defines:

$$M_{\varrho,\boldsymbol{\xi}_0,t_0,\theta} \stackrel{def}{=} \max_{\substack{\boldsymbol{\xi}\in S(\boldsymbol{\xi}_0,\varrho)\\ t\in[t_0-\theta,t_0+\theta]}} |\mathbf{F}(\boldsymbol{\xi},t)| \equiv M \tag{2.3.2}$$

one can choose

$$T_{\varrho,\theta} = \frac{\varrho}{\varrho+\theta M}\,\theta. \tag{2.3.3}$$

Observations.

(1) By Proposition 1, §(2.2), it is enough to show the existence of a $C^{(1)}$ solution verifying Eq. (2.3.1).
(2) The proof that follows is "constructive" in the sense that it provides a sequence $t \to \mathbf{x}^{(n)}(t)$, $t \in [t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$, of functions approximating (as $n \to \infty$) the solution and, at the same time, it provides an estimate of the approximation error defined as $\max|\mathbf{x}(t) - \mathbf{x}^{(n)}(t)|$, where the maximum is taken on the interval $[t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$.
(3) It is often useful, in applications, not to follow the solution scheme proposed by the following proof of Proposition 2. It might, in fact, be more convenient to use *ad hoc* procedures based on the particular features of the $\mathbf{F}$ under analysis in a concrete case. Usually, with such procedures one finds much better error estimates than the ones following from general methods, where one cannot take into account some special properties of the equations (e.g., symmetry properties, Hamiltonian form, etc.).
(4) To understand informally the bound on the magnitude of the interval of existence consider first that, during the proof, it appears necessary to have an a priori control of how far $\mathbf{x}(t)$ can travel away from the initial position $\boldsymbol{\xi}_0$. The continuity of $\mathbf{F}$ guarantees the boundedness of the maximum of $|\mathbf{F}(\boldsymbol{\xi},t)|$, for, say, $\boldsymbol{\xi} \in S(\boldsymbol{\xi}_0,\varrho)$, $t \in [t_0-\theta, t_0+\theta]$. It follows that during the whole time interval $[t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$, the point $\mathbf{x}(t)$ stays inside $S(\boldsymbol{\xi}_0,\varrho)$ because $\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t),t)$ and the right-hand side of this relation does not exceed M, Eq. (2.3.2): notice, in fact, that $T_{\varrho,\theta}$ has been chosen, just to achieve this effect, smaller than both θ and ϱM^{-1} (i.e., $T_{\varrho,\theta} = (\theta^{-1} + \varrho^{-1}M)^{-1}$ so that $MT_{\varrho,\theta} < \varrho$).
(5) The interval $[t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$ is certainly not optimal, at least because the choice of the set $S(\boldsymbol{\xi}_0,\varrho) \times [t_0-\theta, t_0+\theta]$, where the maximum of $|\mathbf{F}|$ is considered, was arbitrary. A better existence interval could be obtained using this arbitrariness and optimizing the result over the possible sets on which one takes the maximum. Also, once the existence of a solution verifying Proposition 2 has been established, one could apply Proposition 2 and Proposition 1 to the equation with initial datum $\mathbf{x}(t_0+T_{\varrho,\theta})$ at the initial time $t_0+T_{\varrho,\theta}$, thus continuing it beyond $T_{\varrho,\theta}$. However one cannot hope, in general, for an infinite existence interval containing $\mathcal{R}_+$: this can be seen through counterexamples. The simplest among them is provided by the equation $\dot{x} = x^2, x(0) = 1$, in $\mathcal{R}$.

PROOF. Rather than studying $C^{(1)}$ solutions of $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$ verifying the initial conditions (2.3.1), look for $\mathcal{R}^d$-valued $C^{(0)}([t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}])$ solutions of the equation:

$$\mathbf{x}(t) = \boldsymbol{\xi}_0 + \int_{t_0}^{t} \mathbf{F}(\mathbf{x}(\tau), \tau)\, d\tau. \tag{2.3.4}$$

Every $C^{(0)}([t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}])$ function verifying Eq. (2.3.4) is a $C^{(1)}$ solution to the original equation also verifying Eq. (2.3.1), and vice versa. For $t \in [t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$ define the sequence of $\mathcal{R}^d$-valued functions $t \to \mathbf{x}^{(n)}(t), n = 0, \ldots$, through the following recursive scheme:

$$\begin{aligned}
\mathbf{x}^{(0)}(t) &= \boldsymbol{\xi}_0, \\
\mathbf{x}^{(1)}(t) &= \boldsymbol{\xi}_0 + \int_{t_0}^{t} \mathbf{F}(\mathbf{x}^{(0)}(\tau), \tau)\, d\tau, \\
&\ldots \\
\mathbf{x}^{(n)}(t) &= \boldsymbol{\xi}_0 + \int_{t_0}^{t} \mathbf{F}(\mathbf{x}^{(n-1)}(\tau), \tau)\, d\tau,
\end{aligned} \tag{2.3.5}$$

and remark that each such function is in $C^{\infty}(\mathcal{R})$ and it s natural to try taking the limit as $n \to +\infty$. The existence, uniformly in $t \in [t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$, of

$$\lim_{n\to\infty} \mathbf{x}^{(n)}(t) = \mathbf{x}(t) \tag{2.3.6}$$

should imply that the limit function will also be continuous. Existence and uniformity of the limit is obtained by rewriting it as

$$\mathbf{x}^{(0)}(t) + \sum_{k=1}^{\infty} (\mathbf{x}^{(k)}(t) - \mathbf{x}^{(k-1)}(t)) \tag{2.3.7}$$

and deducing that if

$$\mu_k = \max_{t\in[t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]} |\mathbf{x}^{(k)}(t) - \mathbf{x}^{(k+1)}(t)|, \qquad \text{then} \tag{2.3.8}$$

$$\sum_{k=0}^{\infty} \mu_k < +\infty \tag{2.3.9}$$

This will mean that the series of Eq. (2.3.7) is uniformly convergent for $t \in [t_0 - T_{\varrho,\theta}, t_0 + T_{\varrho,\theta}]$: hence, the same will hold for the limit of Eq. (2.3.6).

To estimate μ_k we can refer to Eq. (2.3.5) to obtain for $k = 2, 3, \ldots$,

$$\mathbf{x}^{(k)}(t) - \mathbf{x}^{(k-1)}(t) = \int_{t_0}^{t} \big(\mathbf{F}(\mathbf{x}^{(k-1)}(\tau), \tau) - \mathbf{F}(\mathbf{x}^{(k-2)}(\tau), \tau)\big)\, d\tau \tag{2.3.10}$$

Through Lagrange's theorem in the form

$$|\mathbf{F}(\boldsymbol{\xi},\tau)-\mathbf{F}(\boldsymbol{\eta},\tau)|\le L\,|\boldsymbol{\xi}-\boldsymbol{\eta}|,\qquad \forall\,\boldsymbol{\xi},Bh\in S(\boldsymbol{\xi}_0,\varrho),\ \forall\,\tau\in[t_0-T_{\varrho,\theta},t_0+T_{\varrho,\theta}] \tag{2.3.11}$$

where

$$L=\max_{\substack{\boldsymbol{\xi}\in S(\boldsymbol{\xi}_0,\varrho)\\ t\in[t_0-T_{\varrho,\theta},t_0+T_{\varrho,\theta}]}}\sum_{i,j=1}^{d}\Big|\frac{\partial F^{(i)}}{\partial\xi_j}(\boldsymbol{\xi},t)\Big| \tag{2.3.12}$$

Eqs. (2.3.10) and (2.3.11) imply:

$$|\mathbf{x}^{(k)}(t)-\mathbf{x}^{(k-1)}(t)|\le L\int_{[t_0,t]}|\mathbf{x}^{(k-1)}(\tau)-\mathbf{x}^{(k-2)}(\tau)|\,d\tau \tag{2.3.13}$$

$\forall\,k=2,3,\ldots$ provided we preliminarily check that for all $k=0,1,\ldots$, the functions $t-\to\mathbf{x}^{(k)}(t)$, $t\in[t_0-T_{\varrho,\theta},t_0+T_{\varrho,\theta}]$, take their values in $S(\boldsymbol{\xi}_0,\varrho)$.

This last property is proved inductively starting from Eq. (2.3.5): keeping in mind the choice of $T_{\varrho,\theta}$ (chosen, as essentially stated in observation (4), just in such a way to make this property true) suppose, inductively, that $|\mathbf{x}^{(h)}(t)-\boldsymbol{\xi}_0|\le\varrho,\forall\,h=0,\ldots,k-1$; it is a property which holds for $k=1$. To check that $|\mathbf{x}^{(k)}(t)-\boldsymbol{\xi}_0|\le\varrho$ remark that Eqs. (2.3.5) and (2.3.3) give

$$|\mathbf{x}^{(k)}(t)-\boldsymbol{\xi}_0|\le\int_{[t_0,t]}d\tau\,|\mathbf{F}(\mathbf{x}^{(k-1)}(\tau),\tau)|\le M_{\varrho,\boldsymbol{\xi}_0,\theta}\,|t-t_0|<\varrho \tag{2.3.14}$$

Eq. (2.3.13), follows because Eq. (2.3.14) with $k=1$ yields for $t\in[t_0-T_{\varrho,\theta},t_0+T_{\varrho,\theta}]$,

$$|\mathbf{x}^{(k)}(t)-\mathbf{x}^{(k-1)}(t)|\le L^{k-1}\int_{[t_0,t]}d\tau_1\int_{[t_0,\tau_1]}d\tau_2\ldots\times\int_{[t_0,\tau_{k-2}]}d\tau_{k-1}|\mathbf{x}^{(1)}(\tau_{k-2})-\boldsymbol{\xi}_0|\le\frac{L^{k-1}T_{\varrho,\theta}^{k-1}}{(k-1)!}\varrho \tag{2.3.15}$$

since $T_{\varrho,\theta}\ge|t-t_0|$. Eq. (2.3.15) shows the convergence of the series of Eq. (2.3.9) and, therefore, the limit of Eq. (2.3.6) exists uniformly for $t\in[t_0-T_{\varrho,\theta},t_0+T_{\varrho,\theta}]$ and defines a function $t\to\mathbf{x}(t)$ on this interval with values in $S(\boldsymbol{\xi}_0,\varrho)$. It satisfies Eq. (2.3.4) as it is seen by taking the $n\to\infty$ limit in Eq. (2.3.5) and by using the uniformity of the limit of Eq. (2.3.6) to exchange the integration with the limit. mbe

2.3.1 Problems

1. Give a lower estimate for the magnitude of $T_{\varrho,\theta}$, the amplitude of the existence interval as in Proposition 2, for the following second-order equations, assuming $x(0)=0,\dot{x}(0)=1$ or $x(0)=1,\dot{x}(0)=0$ as initial data at $t_0=0$:

$$\ddot{x} = x, \quad \ddot{x} = x + x^3, \quad \ddot{x} = x - \dot{x} + x^3, \quad \ddot{x} = -\dot{x}^2, \quad \ddot{x} == -\sin x.$$

Also estimate $\sup_{\varrho,\theta} T_{\varrho,\theta}$ from below. (*Hint*: Reduce the equation to first order and then apply Proposition 1.)

2. Solve the equation $\ddot{x} = x$ with initial datum $x(0) = 1, \dot{x}(0) = 0$.

3. Solve the equations $\dot{x} = -x^2, \dot{x} = \cos x, \dot{x} = (\cos x)^2$ with initial datum $x(0) = 1$.

4. Solve the equation $\dot{x} = x + y, \dot{y} = -x + 2y$ with initial datum $x(0) = 0, y(0) = 1$.

5. Using the "quadrature method", solve the equation $\ddot{x} = 4(x^3 - x)$, $x(0) = 0, \dot{x} = \sqrt{2}$ (see §2.1, final comment).

6. As in Problem 5 for $\dot{x} = -(4x^3 + 6x^2 - 2), x(0) = 0, \dot{x}(0) = \sqrt{2}$.

7. Find two linearly independent solutions for the equation in Problem 4.

8.* Compute $w(t)$ for the equation in Problem 4 (see Problem 8, §(2.2).

9.* Let $t \to \mathbf{L}(t)$ be a $d \times d$-matrix-valued C^∞ function on $\mathcal{R}$. Show that the equation $\dot{\mathbf{x}}(t) = \mathbf{L}(t)\mathbf{x}(t)$ admits d linearly independent solutions defined for $|t| \le T$ with T small enough. (*Hint*: Let $\mathbf{x}^{(i)}$ be the solution with initial data $x_j^{(i)}(0) = \delta_{i,j}, i, j = 1, \ldots d$. Then evaluate an existence interval for such initial data.)

10.* Compute $T_{1,1}$ for the equation in Problem 9 when $|t_0| < \sigma$ and $\boldsymbol{\xi}_0$ is arbitrary, $\boldsymbol{\xi}_0 = \mathbf{x}(t_0)$; for the symbols, see Proposition 1. Show that $|\boldsymbol{\xi}_0|\, T_{1,1}$ can be taken to be independent of t_0 and $\boldsymbol{\xi}_0$ at a given $\sigma > 0$. Deduce from this that every solution to $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$ can be extended to a solution defined for $t \in \mathcal{R}$.

11. Let $\mathbf{L}$ be a $d \times d$ matrix and consider the equation $\dot{\mathbf{x}} = \mathbf{L}\mathbf{x}$ in $\mathcal{R}^d$. Suppose that $\mathbf{L}$ has d pairwise distinct real eigenvalues (see Appendix E for the eigenvalue notion) $\lambda_1, \ldots, \lambda_d$. Let $\mathbf{v}^{1)}, \ldots, \mathbf{v}^{(d)}$ be the respective real linearly independent eigenvectors (see Appendix E). Show that the functions $t \to e^{\lambda_i t}\mathbf{v}^{(i)}$ are d linearly independent solutions. Show that any solution $t \to \mathbf{x}(t)$ has the form

$$\mathbf{x}(t) = \sum_{j=1}^{d} \alpha_j e^{\lambda_j t} \mathbf{v}^{(j)}, \qquad \text{with} \quad (\alpha_1, \ldots, \alpha_d) \in \mathcal{R}^d.$$

2.4 General Properties of Motion. Regularity.

In proving Proposition 2 it was found that $C^{(1)}$ solutions of $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$, $\mathbf{F} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$, are necessarily C^∞ solutions. This is the simplest regularity property shown by the solutions of such differential equations. Other regularity properties of the solutions will be now analyzed.

In applications it often happens that the right-hand side of Eq. (2.2.4) depends on parameters $\boldsymbol{\alpha} \in \mathcal{R}^m$ and that, furthermore, it is important to know how the solutions change as the initial data $\boldsymbol{\xi}_0$ and the parameters $\boldsymbol{\alpha}$ vary in $\mathcal{R}^d$ and $\mathcal{R}^m$, respectively. A first answer to this question is provided by the following proposition.

3 Proposition. *Let* $\boldsymbol{\xi}, t, \boldsymbol{\alpha} \to \mathbf{F}(\boldsymbol{\xi}, t, \boldsymbol{\alpha})$ *be a* $C^\infty(\mathcal{R}^d \times \mathcal{R} \times \mathcal{R}^m)$ *function taking its values in* $\mathcal{R}^d$*, and consider the equation*

$$\mathbf{x}(t) = \boldsymbol{\xi}_0 + \int_{t_0}^{t} \mathbf{F}(\mathbf{x}(\tau), \tau, \boldsymbol{\alpha}_0)\, d\tau \tag{2.4.1}$$

as an equation for the continuous function $t \to \mathbf{x}(t)$ *parameterized by* $\boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0 \in \mathcal{R}^d \times \mathcal{R} \times \mathcal{R}^m$. *Given* $\varrho, \theta, a > 0$ *and* $(\overline{\boldsymbol{\xi}}, \overline{t}, \overline{\boldsymbol{\alpha}}) \in \mathcal{R}^d \times \mathcal{R} \times \mathcal{R}^m$, *there exists* $T > 0$ *such that:*
(i) Eq. (2.4.1) admits a solution for every $(\boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0)$ *close enough to* $(\overline{\boldsymbol{\xi}}, \overline{t}, \overline{\boldsymbol{\alpha}})$ *such that* $|\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}_0| < \frac{\varrho}{2}$, $|t - t_0| < \frac{\theta}{2}$, $|\overline{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0| < a$. *Such solution will be denoted* $t \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ *and it is defined for* $t \in [t_0 - T, t_0 + T]$.
(ii) The function $S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$, *defined for*

$$|\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}_0| < \frac{\varrho}{2}, \ |t - t_0| < \frac{\theta}{2}, \ |\overline{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0| < a, \ |t - t_0| \le T \tag{2.4.2}$$

takes its values inside the ball $S(\overline{\boldsymbol{\xi}}; \varrho)$ *with center* $\overline{\boldsymbol{\xi}}$ *and radius* ϱ *and it is a* C^∞ *function of its arguments.*
(iii) The value T *can be taken as:*

$$T = \frac{\varrho}{2(\varrho + \theta\, \max |\mathbf{F}(\boldsymbol{\xi}, t, \boldsymbol{\alpha})|)}\, \theta \tag{2.4.3}$$

where the maximum is considered on the set $|\boldsymbol{\xi} - \overline{\boldsymbol{\xi}}| < \frac{\varrho}{2}, |t - \overline{t}| < \frac{\theta}{2}, |\boldsymbol{\alpha} - \overline{\boldsymbol{\alpha}}| < a$.

Observations.

(1) Eq. (2.4.1) is equivalent to

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t, \boldsymbol{\alpha}_0), \qquad \mathbf{x}(t_0) = \boldsymbol{\xi}_0 \tag{2.4.4}$$

and, therefore, the above proposition provides a regularity theorem for the solutions of Eq. (2.4.4) as functions of the initial data, of the initial time, of time itself, and of the parameters $\boldsymbol{\alpha}$ on which $\mathbf{F}$ may possibly depend. The set (2.4.2) and the key estimate (2.4.3) should not be taken too seriously as they are not optimal: they merely show an example of the type of concreteness that can be attained in the formulation of a regularity criterion (see, also, observation 4, p. 19).
(2) Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{d+m+2}) \equiv ((\boldsymbol{\xi}_0)_1, \ldots, (\boldsymbol{\xi})_d, (\boldsymbol{\alpha}_0)_1, \ldots, (\boldsymbol{\alpha}_0)_m, t, t_0)$ and

$$\begin{aligned}\mathbf{x}(t) &= (x_1(t), \ldots, x_d(t)) = S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0) \\ &\equiv (S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)_1, \ldots, (S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)_d)\end{aligned} \tag{2.4.5}$$

Formal differentiation of Eq. (2.4.4) with respect to $\beta_i, i = 1, 2, \ldots, m + d$, gives

$$\frac{d}{dt}\frac{\partial \mathbf{x}(t)}{\partial \beta_i} = \sum_{h=1}^{d} \frac{\partial \mathbf{F}}{\partial \xi_h}(\mathbf{x}(t), t, \boldsymbol{\alpha}_0)\frac{\partial x_h(t)}{\partial \beta_i} + \sum_{h=1}^{d} \frac{\partial \mathbf{F}}{\partial \alpha_k}(\mathbf{x}(t), t, \boldsymbol{\alpha}_0)\frac{\partial \alpha_k}{\partial \beta_i} \tag{2.4.6}$$

$$\Big(\frac{\partial \mathbf{x}(t)}{\partial \beta_i}\Big)_{t=t_0} = \frac{\partial \boldsymbol{\xi}_0}{\partial \beta_i} \tag{2.4.7}$$

Analogous equations for the higher-order derivatives can also obtained.
(3) From the proof that the above $d(m+d)$ derivatives $\frac{\partial x_j(t)}{\partial \beta_i}$ do actually verify these equations.
(4) The $d(m+d)$ equations (2.4.6) and (2.4.7) can be considered by imagining that $\mathbf{x}(t)$ is a known function [obtained by first solving Eq. (2.4.4)]. Then, for each $i = 1, \ldots, m+d$, Eq. (2.4.6) can be thought of as a system of d differential equations for the functions of $t, t \to \frac{\partial \mathbf{x}(t)}{\partial \beta_i}$, with initial data at t_0 given by Eq. (2.4.7). Each such system can be solved by regarding it as an ordinary linear system of differential equations of the type (2.2.4) with suitable initial data. Actually this is a method to compute the derivatives $\frac{\partial x_j(t)}{\partial \beta_i}$ which, as it will appear in several instances, turns out to be quite useful. It is also useful in numerical computations.
(5) Similarly, equations for the t or t_0-derivatives follow from Eq. (2.4.4):

$$\frac{\partial \mathbf{x}(t)}{\partial t} = \mathbf{F}(\mathbf{x}(t), t, \boldsymbol{\xi}_0), \qquad \mathbf{x}(t_0) = \boldsymbol{\xi}_0, \qquad \text{and} \tag{2.4.8}$$

$$\begin{aligned} &\frac{d}{dt}\frac{\partial \mathbf{x}(t)}{\partial t_0} = \sum_{h=1}^{d} \frac{\partial \mathbf{F}(\mathbf{x}(t), t, \boldsymbol{\alpha}_0)}{\partial \xi_h}\frac{\partial x_h(t)}{\partial t_0}, \\ &\Big(\frac{\partial \mathbf{x}(t)}{\partial t_0}\Big)_{t=t_0} = -\mathbf{F}(\boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0) \end{aligned} \tag{2.4.9}$$

to which remarks (3) and (4) apply.
(6) Had $\mathbf{F}$ been assumed to be a $C^{(k)}$ function, $k > 1$, on $\mathcal{R}^d \times R \times \mathcal{R}^m$ one could still have obtained a regularity result: however, one could only show, with the same proof that follows, that the function $(t, \boldsymbol{\xi}_0, \boldsymbol{\alpha}_0, t_0) \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ is a $C^{(k)}$ function in the region of Eq. (2.4.2).
(7) Proposition 3 also yields a regularity theorem for the solutions of higher-order differential equations, of the type considered in Eq. (2.2.2), via the reduction to first order described in Eq. (2.2.3). The explicit statement of the corresponding results is left as a problem for the reader.

PROOF. This proof is essentially a repetition of the proof of Proposition 2, on the existence property. Here a sketch is provided, leaving to the reader the elaboration of the details, if he deems it necessary.

The statement about the existence (and uniqueness) of the solutions $t \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ follows easily from Proposition 2: Proposition 2 also implies the estimate (2.4.3) for T which follows from Eq. (2.3.3), identifying the parameters ϱ, θ of Proposition 2 with $\varrho/2, \theta/2$.

First check that $(\boldsymbol{\xi}_0, t, t_0, \boldsymbol{\alpha}_0) \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ is a $C^{(1)}$ function on the set (2.4.2). Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{m+d+2})$ be defined as in observation 2. As seen in §2.3, $t \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ can be thought of as being obtained via a limit of the functions $t \to \mathbf{x}^{(n)}(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0)$ recursively defined for $t \in [t_0 - T, t_0 + T]$ by

$$\begin{aligned}
&\mathbf{x}^{(0)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)=\boldsymbol{\xi}_0\\
&\dots\\
&\mathbf{x}^{(n)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)=\boldsymbol{\xi}_0+\int_{t_0}^{t}\mathbf{F}(\mathbf{x}^{(n-1)}(\tau,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0),\tau,\boldsymbol{\alpha}_0)\,d\tau
\end{aligned}\tag{2.4.10}$$

for $n=1,2,3,\ldots$. The functions $(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)\to\mathbf{x}^{(n)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)$ are [see Eq. (2.4.10)] C^∞ functions of their arguments, $\forall\, n$. Furthermore, differentiating Eq. (2.4.10) with respect $\beta_i, 1=1,\ldots,m+d+2$, it is:

$$\begin{aligned}
&\frac{\partial\mathbf{x}^{(n)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)}{\partial\beta_i}=\frac{\partial\boldsymbol{\xi}_0}{\partial\beta_i}\\
&+\int_{t_0}^{t}\Big\{\sum_{j=1}^{d}\frac{\partial\mathbf{F}}{\partial\xi_j}(\mathbf{x}^{(n-1)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0),\tau,\boldsymbol{\alpha}_0)\frac{\partial\mathbf{x}_j^{(n-1)}(\tau,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)}{\partial\beta_i}\\
&+\sum_{\ell=1}^{m}\frac{\partial\mathbf{F}}{\partial\xi_\ell}(\mathbf{x}^{(n-1)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0),\tau,\boldsymbol{\alpha}_0)\frac{\partial\alpha_\ell}{\partial\beta_i}\Big\}\,d\tau\\
&+\mathbf{F}(\mathbf{x}^{(n-1)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0),\tau,\boldsymbol{\alpha}_0))\frac{\partial t}{\partial\beta_i}-\mathbf{F}(\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)\frac{\partial t_0}{\partial\beta_i},
\end{aligned}\tag{2.4.11}$$

where the last two terms arise from the contributions from the integration extremes. This relation between $\frac{\partial\mathbf{x}^{(n)}}{\partial\beta_i}$ and $\frac{\partial\mathbf{x}^{(n-1)}}{\partial\beta_i}$ can be used to estimate $\frac{\partial\mathbf{x}^{(n)}}{\partial\beta_i}-\frac{\partial\mathbf{x}^{(n-1)}}{\partial\beta_i}$ along the lines of proof of Proposition 2. By proceeding in the same way and remarking that Eq. (2.4.10), by the choice of T, implies $\forall\, t\in[t_0-T,t_0+T]$ and $\forall\, n=0,1,\ldots$,

$$|\mathbf{x}^{(n)}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)-\boldsymbol{\xi}_0|\le\frac{\varrho}{2},\tag{2.4.12}$$

it follows, from Eqs. (2.4.11) and (2.4.12), existence of two constants $\overline{M},\overline{L}$ [see Eqs. (2.3.12) and (2.3.15)] such that:

$$|\frac{\partial\mathbf{x}^{(n)}}{\partial\beta_i}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)|\le\overline{M}\tag{2.4.13}$$

$$|\frac{\partial\mathbf{x}^{(n)}}{\partial\beta_i}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)-\frac{\partial\mathbf{x}^{(n-1)}}{\partial\beta_i}(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)|\le\frac{\overline{L}^{n-1}}{(n-1)!}\overline{M}\tag{2.4.14}$$

hold for all $(t,\boldsymbol{\xi}_0,t_0,\boldsymbol{\alpha}_0)$ in the region of Eq. (2.4.2) and for all $n=1,2\ldots$.

Then Eqs. (2.4.13) and (2.4.14) imply existence and uniformity, in region of Eq. (2.4.2), of the limit:

$$\varphi_i(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0) = \lim_{n\to\infty} \frac{\partial \mathbf{x}^{(n)}}{\partial \beta_i}(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0)$$
$$\frac{\partial \mathbf{x}^{(0)}}{\partial \beta_i}(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0) + \sum_{n=1}^{\infty} \Big(\frac{\partial \mathbf{x}^{(n)}}{\partial \beta_i}(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0) - \frac{\partial \mathbf{x}^{(n-1)}}{\partial \beta_i}(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0)\Big), \tag{2.4.15}$$

$\forall\, i = 1, 2, \ldots, m + d + 2$. The above limit is, therefore, a continuous function in the region of Eq. (2.4.2).

Since the limit $\lim_{n\to\infty} \mathbf{x}^{(n)}(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0)$ exists and equals $S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$, the uniformity of the limit in Eq. (2.4.15) guarantees permutability of limit and of $\partial/\partial\beta_i$ operations, thereby showing differentiability of $S_t(Bx_0; t_0, \boldsymbol{\alpha}_0)$ in the region of Eq. (2.4.2). It also shows, *en passant*, via the consideration of the limit as $n \to \infty$ of Eq. (2.4.11), the validity of the statements in observation 3.

An essentially identical argument can be developed to show that $(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0) \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ is in class $C^{(p)}$, $\forall\, p \geq 1$, in the region of Eq. (2.4.2). It will suffice to differentiate Eq. (2.4.11) suitably many times to obtain relations analogous to it for the higher derivatives; such relations will then be used to obtain estimates analogous to Eqs. (2.4.13) and (2.4.14). mbe

2.4.1 Exercises and Problems

1. Solve the equation $\ddot{x} - 2\dot{x} + \alpha x = 0, \alpha > 1$, with initial data $x(0) = x_0, \dot{x}(0) = v_0$, by finding two solutions of the form $t \to Ae^{\lambda\, t}$. By taking the limit $\alpha \to 1$ find the solution, with the same initial data, to $\ddot{x} - 2\dot{x} + x = 0$ (using Proposition 3).

2. Show that the equations $\ddot{x} = -\varepsilon x, \varepsilon > 0$, and $\ddot{x} = 0$ have, for the same initial conditions, solutions $x_\varepsilon(t)$ and $x_0(t)$ such that $\lim_{\varepsilon\to 0} x_\varepsilon(t) = x_0(t), \forall\, t \in \mathcal{R}$. However, show that this limit relation is not uniform in $t \in \mathcal{R}$, except for special initial data.

3. Consider the equation $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t, \boldsymbol{\alpha})$ and suppose that $\mathbf{F}(\mathbf{0}, t, \boldsymbol{\alpha}) \equiv \mathbf{0}$. Then, given $R > 0$ and fixed $t_0 = 0$, show the existence of $\varepsilon > 0, \sigma > 0, \varrho > 0$, such that:

$$(1 - \sigma)|\mathbf{w}| \leq |S_t\mathbf{w}| \leq (1 + \sigma)|\mathbf{w}|$$

having denoted $S_t\mathbf{w}$ the solution to the equation with initial datum $\mathbf{w}$ at $t_0 = 0$. (*Hint*: Apply Lagrange's theorem to estimate $|S_t\mathbf{w} - \mathbf{w}|$ in terms of the maximum of $|\mathbf{F}|$ in a suitable set, and then, likewise, $|S_\tau\mathbf{w} - S_\tau\mathbf{0}| \equiv |S_\tau\mathbf{w}|$, (as $S_\tau\mathbf{0} \equiv \mathbf{0}$), for $|t| \leq \varepsilon, |\mathbf{w}| \leq \varrho$: use the regularity theorem to bound the derivatives of $t, \mathbf{w}, \boldsymbol{\alpha} \to S_\tau\mathbf{w}$; see observations 2-5 to Proposition 3.)

2.5 Local and Global Solutions of Differential Equations

The theory developed so far for the equation:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \qquad \mathbf{x}(t_0) = \boldsymbol{\xi}_0, \tag{2.5.1}$$

where $\mathbf{F}$ is an $\mathcal{R}^d$-valued function in $C^\infty(\mathcal{R}^d \times \mathcal{R})$, is a "local theory"; the existence theorem given in Proposition 2, §2.3, gives, in fact, a solution to Eq. (2.5.1) defined in a finite neighborhood of t_0. It is often necessary in applications to have "global solutions", i.e., solutions to Eq. (2.5.1) defined in time intervals containing a neighborhood of $\mathcal{R}^+ = [0, +\infty)$. In analyzing this problem, the following definition is useful.

2 Definition. *A solution $t \to S_t(\boldsymbol{\xi}_0; t_0)$ of Eq. (2.5.1) defined for $t \in (a, b)$ is called "maximal" if there are no other solutions defined in open intervals properly containing (a, b).*
Two solutions of Eq. (2.5.1) defined in two open intervals I_1 and I_2 coincide in $I_1 \cap I_2$ (see Proposition 1, p. 14); if $I_2 \supset I_1$, the second solution is said a "continuation" of the first.

Observations.

(1) A solution to Eq. (2.5.1) is, therefore, maximal if and only if it "cannot be continued".
(2) For every initial datum $\boldsymbol{\xi}_0$ and every initial time t_0, there is a solution to Eq(2.5.1) which is maximal: the interval of definition of such a solution is the union of all open intervals on which it is possible to define a solution.
(3) This maximality definition only involves open intervals; however, this notion would be the same even other types of intervals were allowed in the definition of maximality. To understand this, just use the existence theorem of §(2.3), p. 18, to continue solutions out of closed or half-closed intervals.

The following proposition clarifies the above notion by showing that a solution of a differential equation can be non global in the future (or in the past) if and only if it "diverges in a finite time".

4 Proposition. *Let $t \to S_t(\boldsymbol{\xi}_0; t_0)$ be a maximal solution for Eq. (2.5.1) and (a, b) be the interval on which this solution is defined. If $b < +\infty$, it must be*

$$\limsup_{t \to b^-} |S_t(\boldsymbol{\xi}_0; t_0)| = +\infty; \tag{2.5.2}$$

if $a > -\infty$, it must be

$$\limsup_{t \to a^+} |S_t(\boldsymbol{\xi}_0; t_0)| = +\infty. \tag{2.5.3}$$

PROOF. Assume $b < +\infty$ and that Eq. (2.5.2) does not hold. Then there exists $K < +\infty$ such that

$$|S_t(\boldsymbol{\xi}_0; t_0)| \le K, \qquad \forall\, t \in [t_0, b) \tag{2.5.4}$$

Using Proposition 2, we can find for every $\tau \in [t_0, b)$ a solution to the equation:

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}(t), t), \qquad \mathbf{x}(\tau) = S_\tau(\boldsymbol{\xi}_0; t_0) \tag{2.5.5}$$

defined for $t \in [\tau - T_1, , \tau + T_1]$, where T_1, by Eq. (2.3.3) with $\varrho = \theta = 1$, can be chosen as

$$T_1 = \big(1 + \max_{\substack{|t-\tau|\le 1 \\ |\boldsymbol{\xi} - \mathbf{x}(\tau)|\le 1}} |\mathbf{F}(\boldsymbol{\xi}, t)|\big)^{-1} \ge \big(1 + \max_{\substack{|\tau|\le 1+|a|+|b| \\ |\boldsymbol{\xi}|\le 1+K}} |\mathbf{F}(\boldsymbol{\xi}, t)|\big)^{-1} \equiv \overline{T}_1) \quad (2.5.6)$$

The solution under investigation can therefore be extended to a solution defined for

$$t \in \cup_{\tau \in (t_0, b)} \big(\tau - \overline{T}_1, \tau + \overline{T}_1\big), \quad (2.5.7)$$

manifestly contradicting the supposed maximality of (a, b). A similar argument holds if $a > -\infty$. mbe

Considering Proposition 4, it is convenient to introduce the following definition.

3 Definition. *Consider the differential equation* $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$ *with* $\mathbf{F}$ *being an* $\mathcal{R}^d$*-valued* $C^\infty(\mathcal{R}^d \times \mathcal{R})$ *function. Suppose that there is an* $\mathcal{R}_+$*-valued continuous function defined on* $\mathcal{R}^3$*:* $(r, s, t) \to \mu(r, s, t)$ *such that if* $t \to S_t(\boldsymbol{\xi}_0; t_0)$ *is a solution to Eq. (2.5.1) defined for* $t \in (a, b)$ *then:*

$$|S_t(\boldsymbol{\xi}_0; t_0)| \le \mu(r, t_0, t), \qquad \forall\ |\boldsymbol{\xi}_0| \le r,\ t_0 \le t. \quad (2.5.8)$$

The differential equation is said "normal" in the future if μ *can be chosen to be* (r, t)*-independent. If* μ *is bounded as* $t \to +\infty$ *the equation will be said to have "bounded trajectories" in the future.*

Observations.

(1) Eq. (2.5.8) is a strong condition on the motions generated by $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$; because of its independence on the existence interval(a, b), it is often called an "a priori estimate" on the motions governed by $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$.
(2) An equation of higher order, like Eq. (2.2.2), will be called normal, or with bounded trajectories, if once reduced to a first-order equation it becomes a normal equation, or an equation with bounded trajectories, in the sense just introduced. More concretely, this means that it is possible to give an a priori estimate (i.e., independent of the interval of definition) of the sizes of $\mathbf{x}(t), \dot{\mathbf{x}}(t), \ldots, \frac{d^{s-1}\mathbf{x}}{dt^s}(t)$ in terms of the observation time t, of the initial time t_0, and of the initial data $\mathbf{x}(t_0), \dot{\mathbf{x}}(t_0), \ldots, \frac{d^{s-1}\mathbf{x}}{dt^s}(t_0)$; furthermore, the bound depends continuously on those parameters.

The importance of the definition is manifest in the following proposition.

5 Proposition. *If the differential equation (2.5.1),* $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$*, is normal, then it admits a "global solution", i.e., a solution defined in a neighborhood of* $[t_0, +\infty)$*, for any given initial datum* $\boldsymbol{\xi}_0$ *and initial time* t_0*.*

PROOF. Let (a,b) be a maximal existence interval for a solution to Eq. (2.5.1), and suppose that $b < +\infty$. Then by Definition 3:

$$\limsup_{t\to b^-} |S_t(\boldsymbol{\xi}_0;t_0)| \le \mu(|\boldsymbol{\xi}_0|,t_0,b) \tag{2.5.9}$$

would hold, contradicting Proposition 4, Eq. (2.5.2). mbe

An example of a normal equation (which also has bounded trajectories) is provided by the following proposition.

6 Proposition. *Consider the differential equation $m\ddot{x} = f(x)$ in $\mathcal{R}$, [see Eq. (2.1.1)], describing the motions of a point with mass $m > 0$, on a line and subject to a force depending only on the position, $f \in C^\infty(\mathcal{R})$. Suppose that the potential energy V, see Eq. (2.1.3), is bounded below. Then the differential equation is normal. If $\lim_{x\to\pm\infty} V(x) = +\infty$, the differential equation also has bounded trajectories.*

PROOF. If $t \to x(t)$ is a solution to Eq. (2.1.1), with $x(t_0) = \xi_0, \dot{x}(t_0) = \eta_0$ and defined for $t \in (a,b)$, by energy conservation (see §(2.1):

$$\frac{1}{2}m\dot{x}^2 + V(x(t)) = E = \frac{1}{2}m\dot{\eta}^2 + V(\xi), \qquad \forall\, t \in (a,b), \tag{2.5.10}$$

and therefore, if $M = \inf_{\xi\in R} V(\xi)$:

$$|\dot{x}(t)| = \sqrt{\frac{2}{m}(E - V(x(t)))} \le \big(\frac{2}{m}(E-M)\big)^{\frac{1}{2}} \tag{2.5.11}$$

and $M > -\infty$, by assumption. Furthermore,

$$|x(t)| = |\xi_0 + \int_{t_0}^t \dot{x}(\tau)\,d\tau| \le |\xi_0| + \big(\frac{2}{m}(E-M)\big)^{\frac{1}{2}}|t-t_0| \tag{2.5.12}$$

which, calling $\mu(|\xi_0|,t_0,t)$ the right-hand side of Eq. (2.5.12), yields an a priori estimate, showing normality.

If $\lim_{\xi\to\pm\infty} V(\xi) = +\infty$, let $\xi \to W(\xi)$ be a symmetric (i.e., $W(\xi) \equiv W(-\xi)$) continuous function which is strictly increasing for $\xi > 0$ and which is a lower bound to $V(\xi): V(\xi) \ge W(\xi), \forall\, \xi \in \mathcal{R}$, and such that $\lim_{\xi\to\infty} W(\xi) = +\infty$. Since V is supposed bounded below, such a function does exist.

Let $\mu(E)$ be the positive solution to $W(\xi) = E$, existing for all $E > M$, i.e. for all E's of the form (2.5.10). Then the motion with energy E given by by th right-hand side of Eq. (2.5.10) must verify $|x(t)| \le \mu(E)$, as $|x(t)| > \mu(E)$ would imply, by the left-hand side of Eq. (2.5.10) and by the choice of W, that $\frac{1}{2}m\dot{x}(t)^2 < 0$.

By the assumed continuity and strict monotonicity of W, the function $\mu(E)$ is continuous in E for $E > M$; hence, (t_0,t)-independent a priori bound $|x(t)| < \mu(E)$ has been obtained. mbe

This section will be concluded by the following remark "a priori estimates". In applications one often meets functions $(\mathbf{x}, t) \to \mathbf{F}(\mathbf{x}, t)$ which are C^∞ functions for $(\mathbf{x}, t) \in (\mathcal{R}^d/A) \times \mathcal{R})$, where A is a "singularity set" usually consisting of points, lines, surfaces, or even in a set with interior points; inside $A \times \mathcal{R}$ the $\mathbf{F}$ might be undefined. In such cases the singularity of $\mathbf{F}$ means that the model originating the differential equations (2.5.1) is not a good model of the physical phenomenon that it hopes to describe, at least if the initial data or the motion generated by them enter the region A.

For instance, the attractive force exerted by the Sun on the Earth is well described by the formula $k|\mathbf{x}|^{-2}$ only if the distance between the Earth and the Sun is large compared to the Sun diameter; it is clear that the singularity in $\mathbf{x} = \mathbf{0}$ is purely fictitious and due to an excessive idealization!

In such cases one is free to modify $\mathbf{F}$ by changing it into a function $\mathbf{F}^{(A)} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$ which, outside a small neighborhood of $A \times \mathcal{R}$, coincides with $\mathbf{F}$. The equation

$$\dot{\mathbf{x}} = \mathbf{F}^{(A)}(\mathbf{x}, t) \tag{2.5.13}$$

will then be an equally good model of the same physical phenomenon.

However, it is obvious that the only interesting motions, among those described by Eq. (2.5.13), will be those evolving outside a neighborhood of $A \times \mathcal{R}$, where, in fact, $\mathbf{F}$ and $\mathbf{F}^{(A)}$ are indistinguishable.

In this book equations of the form (2.5.1) with $\mathbf{F}$ singular in some region will occasionally be considered. However, in all those cases it will also be possible to establish an "a priori estimate" guaranteeing the existence of a continuous positive function μ' on $\mathcal{R}^d \times \mathcal{R} \times \mathcal{R}$ such that if $S_t(\boldsymbol{\xi}_0; t_0)$ is a solution to Eq. (2.5.13), defined for $t \in (a, b)$ with initial datum at t_0 given by $\boldsymbol{\xi}_0 \in \{$set of initial data "thought of as interesting"$\} = \widetilde{A}$, then

$$d(S_t(\boldsymbol{\xi}_0; t_0), \widetilde{A}) \geq \mu'(\boldsymbol{\xi}_0, t, t_0) \tag{2.5.14}$$

where $d(\boldsymbol{\xi}, \widetilde{A}) =$ (distance of $\boldsymbol{\xi}$ from $\widetilde{A}$) and μ' is positive for $\boldsymbol{\xi}_0 \in \widetilde{A}$. Usually one shall fix $\widetilde{A} = A^c =$ (complement of A) by possibly enlarging the set A.

By what has been said so far, it appears that if we are interested only in motions starting outside A and $\widetilde{A} = A^c$, we shall imagine that such motions verify Eq. (2.5.13) and, therefore, we shall be able to apply to them the various results concerning the differential equations with right-hand side in C^∞.

The above elucubrations motivate the following definition:

4 Definition. *Let* $(\boldsymbol{\xi}, t) \to \mathbf{F}(\boldsymbol{\xi}, t)$ *be a* C^∞ *function defined on* $(\mathcal{R}^d/A) \times \mathcal{R}$ *with values in* $\mathcal{R}^d$*, where* $A \subset \mathcal{R}^d$ *is a closed set. Suppose that:*
(i) there exists an $\mathcal{R}^d$*-valued function* $\mathbf{F}^{(A)} \in C^\infty(\mathcal{R}^d \times \mathcal{R} \times \mathcal{R}^d)$ *coinciding with* $\mathbf{F}$ *on* $(\mathcal{R}^d/A) \times \mathcal{R}$*;*
(ii) there exists a real valued function μ' *on* $\mathcal{R}^d \times \mathcal{R} \times \mathcal{R}$*, continuous and positive valued on* $(\mathcal{R}^d/A) \times \mathcal{R} \times \mathcal{R}$*, such that* $t \to \mathbf{x}(t)$ *is a motion verifying* $\dot{\mathbf{x}}(t_0) = \boldsymbol{\xi}_0, \dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{x}(t), t), \forall\, t \in (a, b)$*, then* $\forall\, t \geq t_0$ *and* $t \in (a, b)$:

$$d(S_t(\boldsymbol{\xi}_0, t_0), A) > \mu'(\boldsymbol{\xi}_0, t, t_0) > 0; \tag{2.5.15}$$

(iii) the differential equation $\dot{\mathbf{x}} = \mathbf{F}^{(A)}(\mathbf{x}, t)$ *is normal.*

In such a situation we shall say that the "singular differential equation" $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$ *is "normal outside* A*".*

It is an exercise to prove the following proposition.

7 Proposition. *Let* $(\boldsymbol{\xi}, t) \to \mathbf{F}(\boldsymbol{\xi}, t)$ *be an* $\mathcal{R}^d$*-valued* C^∞ *function on* $(\mathcal{R}^d/A) \times \mathcal{R}$ *and consider the singular differential equation* $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$*: if this equation is normal outside* A*, every initial datum* $\boldsymbol{\xi}_0 \notin A$ *originates a* C^∞ *solution of*

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t); \quad \mathbf{x}(t_0) = \boldsymbol{\xi}_0; \quad \mathbf{x}(t) \notin A \tag{2.5.16}$$

defined in a neighborhood of $[t_0, +\infty)$*, i.e., a global solution.*

Observation. As the reader will verify when looking at Chapter 4, §4.8, §4.9, an interesting example of the situation contemplated in Definition 4 and Proposition 7 can be found in the two-body problem: the set A will be, in this case, the closure of a neighborhood of the set of the initial data with vanishing areal velocity. Such data are those in which the two bodies are heading into or out of a collision and which are, therefore, to be considered singular.

2.5.1 Exercises and Problems

1. Formulate the notions of normal differential equation "in the past" or of differential equation with bounded trajectories "in the past", and reformulate all the propositions of §2.5 to deal with the problem of the existence of solutions in intervals like $t \in (-\infty, t_0]$ or $t \in (-\infty, +\infty)$.

2. Consider the equation in $\mathcal{R}$, $\ddot{x} + \frac{d}{dx}\log(1 + x^2) = 0$. Determine whether it is normal and with bounded trajectories. Compute $x(1)$ with a 60% approximation if $x(0) = 0, \dot{x}(0) = 1$.

3. Same as Problem 2 for $\ddot{x} + \sin x = 0, x(0) = 0, \dot{x}(0) = \frac{1}{4}$.

4. Same as Problems 2 and 3 for the differential equations in Problems 1 and 2 of §2.3.

5.* Same as Problem 2 but with a 1% approximation and using a desk computer together with the error estimate implicit in the existence theorem of §2.3. Alternatively, use the algorithm of Appendix O, together with a desk computer.

6.* Same as Problem 5 but using energy conservation and the relative quadrature formula, together with a desk computer.

7.* Same as Problem 6 but for the equation in Problem 3.

8. Let $t \to \mathbf{x}(t)$ be an $\mathcal{R}^d$-valued $C^\infty(\mathcal{R})$ function such that $\exists M > 0$ for which

$$|\mathbf{x}(t)| \le |\mathbf{x}(0)| + M \int_0^t |\mathbf{x}(\tau)|\, d\tau, \qquad t \ge 0.$$

Show that $|x(t)| \le y(t)$, $t \ge 0$, where y is defined as the solution of $y(t) = |\mathbf{x}(0)| + M \int_0^t y(\tau)\, d\tau \ge 0$, i.e., $y(t) = |\mathbf{x}(0)|e^{Mt}$.

9.* If $\xi \to \varphi(\xi)$ is a continuous positive monotonically increasing function of $\xi \in \mathcal{R}^+$ and if $t \to \mathbf{x}(t)$ is in $C^\infty(\mathcal{R})$ and

$$|\mathbf{x}(y)| \le |\mathbf{x}(0)| + \int_0^t \varphi(|\mathbf{x}(\tau)|)\, d\tau, \qquad t \ge 0$$

show that $|\mathbf{x}(t)| \le y(t)$, $t \ge 0$, where y is defined as the solution of $y(t) = |x(0)| + \int_0^t \varphi(y(\tau))\, d\tau$, i.e., setting $\Phi(y) = \int_{|\mathbf{x}(0)|}^y \frac{\varphi(\eta)}{d\eta}$, as the function verifying $\Phi(y(t)) \equiv t$ (or $y(t) \equiv \Phi^{-1}(t)$).

10.* Given the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ in $\mathcal{R}^d$, define for $T > 0$: $\varphi_T(s) \overset{def}{=} \max_{\substack{t\in[0,T] \\ |\boldsymbol{\xi}|\le s}} |\mathbf{f}(\boldsymbol{\xi}, t)|$
Show that a sufficient condition for the normality of the equation is that, in $\mathcal{R}$,

$$\dot{y} = \varphi_T(y), \qquad y(0) = |\mathbf{x}(0)|$$

admits a global solution (i.e., a solution on $[0, +\infty)$) for all $T > 0$. (*Hint:* $\mathbf{x}(t) = \mathbf{x}(0) + \int_0^t f(x(\tau), \tau)\, d\tau \Rightarrow |\mathbf{x}(t)| \le |\mathbf{x}(0)| + \int_0^t \varphi_T(|\mathbf{x}(\tau)|)\, d\tau$; then apply Problem 9.)

11.* If $t \to \mathbf{L}(t)$ is a matrix-valued $C^\infty(\mathcal{R})$ function with values in the $d \times d$ matrices, the equation $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x}$ is normal in the future (as well as in the past); hence, it has global solutions (*Hint*: Apply Problem 10.)

12. In the context of Problem 11, show that the equation admits d linearly independent global solutions (defined on $(-\infty, +\infty)$). (*Hint*: Use Problem 11 and Problem 9 of §2.3).

13. In the context of Problem 11, suppose that $\mathbf{L}(t)$ is a time-independent matrix $\mathbf{L}$. Using the results of Problem 11 of §2.3, p. 22, and supposing that all the eigenvalues of $\mathbf{L}$ are real and pairwise distinct, show that the equation $\dot{\mathbf{x}} = \mathbf{L}\mathbf{x}$ has bounded trajectories if and only if $\lambda_j \le 0$, $j = 1, \ldots, d$.

14.* In the context of Problem 11, let $\mathbf{g} \in C^\infty(\mathcal{R})$ be an $\mathcal{R}^d$-valued function. Show the normality of the equation $\dot{\mathbf{x}} = \mathbf{L}(t)\mathbf{x} + \mathbf{g}(t)$.

15.* Consider a differential equation in $\mathcal{R}^d$, $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$, with $\mathbf{F} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$. Suppose that $|\mathbf{F}(\mathbf{x}, t)| \le \gamma(t)|\mathbf{x}| + \beta(t)$, where $\beta, \gamma \in C^\infty(\mathcal{R})$, $\beta, \gamma \ge 0$. Show that the equation is normal by finding an a priori estimate. (*Hint*: Combine Problems 9 and 4.)

16.* Same as Problem 5 with $|\mathbf{F}(\mathbf{x}, t)| \le \beta(t) + \gamma(t) \log(e + |\mathbf{x}|)$.

17. Consider a differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t, \boldsymbol{\alpha})$ of the type considered in §2.4, $\mathbf{f} \in C^\infty(\mathcal{R}^d \times \mathcal{R} \times \mathcal{R}^m)$. Suppose that this equation admits an a priori estimate like Eq. (2.5.8), for $\forall\, \boldsymbol{\alpha} \in \mathcal{R}^m$, with an $\boldsymbol{\alpha}$-independent function. Show that, in this case, the "local regularity theorem", Proposition 3, p. 22, becomes "global", i.e., the function $(t, \boldsymbol{\xi}_0, t_0, \boldsymbol{\alpha}_0) \to S_t(\boldsymbol{\xi}_0; t_0, \boldsymbol{\alpha}_0)$ is a C^∞-function of $\boldsymbol{\xi}_0 \in \mathcal{R}^d$, $\boldsymbol{\alpha}_0 \in \mathcal{R}^m$, $t_0 \in \mathcal{R}, t \in \mathcal{R}$, $t \ge t_0$.

2.6 More on Differential Equations. Autonomous Equations

Before proceeding in the analysis of some applications, it is convenient to set up a few more definitions, mainly as an excuse to illustrate some simple but interesting general remarks about differential equations.

5 Definition. *Let* $(\boldsymbol{\xi}, \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(s-1)}) \to \mathbf{f}(\boldsymbol{\xi}, \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(s-1)})$ *be an* $\mathcal{R}^d$ *valued* C^∞ *function on* $\mathcal{R}^{sd}$. *Consider the equation for the* $\mathcal{R}^d$*-valued function* $t \to \mathbf{x}(t)$ *defined for* t *in an interval* I *[see Eq. (2.2.2)]:*

$$\frac{d^s\mathbf{x}}{dt^s} = \mathbf{f}(\mathbf{x}, \frac{d\mathbf{x}}{dt}, \ldots, \frac{d^{s-1}\mathbf{x}}{dt^{s-1}}), \tag{2.6.1}$$

$$\mathbf{x}(t_0) = \boldsymbol{\xi}_0, \ \frac{d\mathbf{x}}{dt}(t_0) = \boldsymbol{\xi}^{(1)}, \ldots, \frac{d^{s-1}\mathbf{x}}{dt^{s-1}}(t_0) = \boldsymbol{\xi}^{(s-1)}. \tag{2.6.2}$$

Eq. (2.6.1) will be called an "autonomous" differential equation of class C^∞. *In other words, Eq. (2.2.2) is said to be autonomous when the right-hand side "does not explicitly depend upon time".*
The space $\mathcal{R}^d \times \ldots \times \mathcal{R}^d = \mathcal{R}^{sd}$, *thought of as the space of the possible initial data* $(\boldsymbol{\xi}, \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(s-1)})$ *for Eq. (2.6.1), will be called the "space of the initial data" or the "data space".*

It is also useful to introduce the following definition.

6 Definition. *We shall say that a* C^∞ *autonomous differential equation like Eq. (2.6.1) is "reversible" if any solution,* $t \to \mathbf{x}(t)$, *to Eq. (2.6.1) defined for* $t \in (-\varepsilon, \varepsilon)$, $\varepsilon > 0$, *is such that the function* $t \to \mathbf{x}(-t)$, $t \in (-\varepsilon, \varepsilon)$, *is also a solution to Eq. (2.6.1).*

Observation. We shall see that many differential equations describing non dissipative dynamical systems are reversible. Basically, $\mathbf{f}$ originates a reversible system when s is even and $\mathbf{f}$ depends evenly on the odd derivatives. It should be kept in mind that $t \to \mathbf{x}(-t)$ will in general be a solution which corresponds to different initial data (unless $s = 1$): for instance $\ddot{x} = x$ is an equation in $\mathcal{R}^2$ which is reversible, but its solution $x(t) = e^t$ has initial data $x(0) = 1, \dot{x}(0) = 1$ while the solution $t \to e^{-t}$ has initial data $x(0) = 1, \dot{x}(0) = -1$.

The interest in autonomous equations lies, from a mathematical point of view, in the validity of the following easy propositions.

8 Proposition. *Consider a normal autonomous first-order.*[4] *differential equation in* $\mathcal{R}^d$. *It is possible to define on* $\mathcal{R}^d$ *a family* $(S_t)_{t\geq 0}$ *of maps, mapping* $\mathcal{R}^d$ *into itself, such that the functions*

$$t \to S_{t-t_0}(\boldsymbol{\xi}_0), \qquad \boldsymbol{\xi}_0 \in \mathcal{R}^d, \ t, t_0 \in \mathcal{R}, \ t \geq t_0 \tag{2.6.3}$$

solve Eq. (2.6.1) with initial datum at $t = t_0$ *given by* $\boldsymbol{\xi}_0$. *For every* $t \geq 0$, *the map* S_t *is a* C^∞ *map and*

$$S_t(S_{t'}(\boldsymbol{\xi})) = S_{t+t'}(\boldsymbol{\xi}), \qquad \forall\, t, t' \geq 0, \ \forall\, \boldsymbol{\xi} \in \mathcal{R}^d. \tag{2.6.4}$$

Furthermore, the maps S_t *are* C^∞ *regular jointly in* t_0 *and* t*: i.e., the functions* $(t, \boldsymbol{\xi}_0) \to S_t(\boldsymbol{\xi}_0)$, $(t, \boldsymbol{\xi}_0) \in \mathcal{R}_+ \times \mathcal{R}^d$ *are in* $C^\infty(\mathcal{R}_+ \times \mathcal{R}^d)$.

[4] I.e., $s = 1$ in Eq. (2.6.1)

PROOF. Let $t \to S_t(\boldsymbol{\xi}_0; t_0)$ be the solution to Eqs. (2.6.1) and (2.6.2) with $s = 1$, defined for $t > t_0$. Such a solution does exist since Eq. (2.6.1) is now supposed to be a normal equation. Let:

$$S_t(\boldsymbol{\xi}_0) = S_t(\boldsymbol{\xi}_0; 0) \qquad \text{for } t \geq 0 \tag{2.6.5}$$

From §2.5, S_t is a C^∞ map of $\mathcal{R}^d$ into itself for each $t \in \mathcal{R}_+$ and, also, that $(t, \boldsymbol{\xi}_0) \to S_t(\boldsymbol{\xi}_0)$, $(t, \boldsymbol{\xi}_0) \in \mathcal{R}_+ \times \mathcal{R}^d$ is in $C^\infty(\mathcal{R}_+ \times \mathcal{R}^d)$. For $t \geq t_0$, let $\mathbf{x}(t) = S_{t-t_0}(\boldsymbol{\xi}_0)$. Since $\mathbf{f}$ "does not explicitly depend on time", it is

$$\begin{aligned} \frac{d\mathbf{x}}{dt}(t) &= \frac{d}{dt} S_{t-t_0}(\boldsymbol{\xi}_0) = \frac{d}{dt} S_{t-t_0}(\boldsymbol{\xi}_0, 0) \\ &= \mathbf{f}(S_{t-t_0}(\boldsymbol{\xi}_0, 0)) = \mathbf{f}(S_{t-t_0}(\boldsymbol{\xi}_0)) = \mathbf{f}(\mathbf{x}(t)) \end{aligned} \tag{2.6.6}$$

Hence $t \to S_{t-t_0}(\boldsymbol{\xi}_0)$ is a solution to Eq. (2.6.1) for $t \geq t_0$. Furthermore,

$$S_{t_0-t_0}(\boldsymbol{\xi}_0) = S_0(\boldsymbol{\xi}_0) = S_0(\boldsymbol{\xi}_0; 0) \equiv \boldsymbol{\xi}_0 \tag{2.6.7}$$

which, by the uniqueness theorem, Proposition 1, p.14, gives $S_{t-t_0}(\boldsymbol{\xi}_0) \equiv S_t(\boldsymbol{\xi}_0, t_0)$, $t \geq t_0$ Similarly, one checks that $t \to S_t(S_{t'}(\boldsymbol{\xi}_0))$ is a solution to Eq. (2.6.1) with initial datum at $t = 0$ equal to $S_{t'}(\boldsymbol{\xi}_0)$; such is also $\to S_{t+t'}(\boldsymbol{\xi}_0)$; hence, Eq. (2.6.4) is also proved. mbe

9 Corollary. *Consider an autonomous equation of order s, as in Eq. (2.6.1), and suppose that it is normal. It is possible to define, on the data space $\mathcal{R}^d$, a family $(S_t)_{t\geq 0}$ of C^∞ maps of $\mathcal{R}^{ds}$ into itself such that the function*

$$t \to S_{t-t_0}(\boldsymbol{\xi}_0, \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(s-1)}) = (\mathbf{x}(t), \mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(s-1)}(t)) \tag{2.6.8}$$

is a solution to the equations

$$\begin{aligned} &\dot{\mathbf{x}}(t) = \mathbf{x}^{(1)}(t),\ \dot{\mathbf{x}}^{(1)}(t) = \mathbf{x}^{(2)}(t), \ldots, \dot{\mathbf{x}}^{(s-2)}(t) = \mathbf{x}^{(s-1)}(t), \\ &\dot{\mathbf{x}}^{(s-1)}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(s-1)}(t)) \end{aligned} \tag{2.6.9}$$

[*equivalent to Eq. (2.6.1)*] *and verifies the initial data*

$$\mathbf{x}(t_0) = \boldsymbol{\xi}_0,\ \mathbf{x}^{(1)}(t_0) = \boldsymbol{\xi}^{(1)}, \ldots,\ \mathbf{x}^{(s-1)}(t_0) = \boldsymbol{\xi}^{(s-1)}. \tag{2.6.10}$$

Furthermore,

$$S_t S_{t'} = S_{t+t'}, \qquad \forall\, t, t' \geq 0 \tag{2.6.11}$$

and the maps S_t are C^∞ regular also, jointly in t and $(\boldsymbol{\xi}_0, \ldots, Bx^{(s-1)})$; i.e., the map $(t, \boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{s-1}) \to S_t(\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{s-1})$, with $(t, \boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{s-1}) \in \mathcal{R}_+ \times \mathcal{R}^d \times \ldots \mathcal{R}^d$ is in $C^\infty(\mathcal{R}_+ \times \mathcal{R}^d \times \ldots \mathcal{R}^d)$.

PROOF. It is an immediate consequence of the equivalence between Eqs. (2.6.1), (2.6.2) and Eqs. (2.6.9), (2.6.10) and of Proposition 8.

mbe

7 Definition. *Given a normal s-th order autonomous differential equation on $\mathcal{R}^d$, the family $(S_t)_{t\geq 0}$ of maps of the data space into itself, defined in Proposition 8, will be called the "flow" on $\mathcal{R}^d$ which "solves Eqs. (2.6.1) and (2.6.2)".*

Observations.

(1) Because of Eq. (2.6.11), the flow $(S_t)_{t\geq 0}$ is, in mathematical language, a "semigroup". When Eq. (2.6.1) is also normal in the past, it becomes possible to define S_t for $t \leq 0$, and the family $(S_t)_{t\in\mathcal{R}}$ forms a group, i.e., it verifies Eq. (2.6.11) for all $t, t' \in \mathcal{R}$ (exercise).
(2) All the normal reversible equations are also normal in the past (exercise); hence, such a class of equations provides an important instance when the solution flow is a group.

An interesting remark about autonomous equations, already met in Problem 3, §2.2, is the following proposition.

10 Proposition. *Consider a normal s-th order autonomous differential equation on $\mathcal{R}^d$, like Eq. (2.6.1). Suppose that $(\boldsymbol{\xi}_0^{(0)}, \boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$ is an initial datum such that there is some $T > 0$ for which*

$$S_T(\boldsymbol{\xi}_0^{(0)}, \boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)}) = (\boldsymbol{\xi}_0^{(0)}, \boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)}); \qquad (2.6.12)$$

then the motion generated by $(\boldsymbol{\xi}_0^{(0)}, \boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$ is a "periodic motion" with period T, i.e., it is a periodic solution of Eq. (2.6.1) with period T.

PROOF. The function $t \to S_{t+T}(\boldsymbol{\xi}_0^{(0)}(\boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$, $t > 0$, where $(S_t)_{t\geq 0}$ is the solution flow to Eq. (2.6.1), is again a solution to Eq. (2.6.1) and, for $t = 0$, verifies the initial condition $(\boldsymbol{\xi}_0^{(0)}, \boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$ by our assumption Eq. (2.6.12). Hence, by uniqueness, it coincides with $t \to S_t(\boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$. This means that $t \to S_t(\boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$ is periodic with period T. mbe

Observation. More generally, it is clear that $t \to S_{t+T}(\boldsymbol{\xi}_0^{(1)}, \ldots, \boldsymbol{\xi}_0^{(s-1)})$ is a solution to Eq. (2.6.9) for $t \geq 0$: i.e., if $t \to \mathbf{x}(t)$ is a solution to an autonomous equation, $t \to \mathbf{x}(t+T)$ is also a solution for $T \in \mathcal{R}$.

2.6.1 Exercises and Problems

Show that the following equations are normal both in the past and in the future and:

1. Draw the trajectories of the flow $(S_t)_{t\in\mathcal{R}}$ in the data space $\mathcal{R}^2$ for the equation $\ddot{x} = -g$, $g \in \mathcal{R}$.

2. Same as Problem 1 for $\ddot{x} = -\omega^2 x$, $\omega^2 > 0$

3. Same as Problem 1 for $\ddot{x} = -g - \lambda\dot{x}$, $g, \lambda \in \mathcal{R}$.

4, Same as Problem 1 for $\ddot{x} = \omega^2 x$, $\omega^2 > 0$.

5. Describe the trajectories of the flow $(S_t)_{t\in\mathcal{R}}$ in the data space $\mathcal{R}^6$ for the equations in $\mathcal{R}^3$: $\ddot{\mathbf{x}} = -\mathbf{g}$ or $\ddot{\mathbf{x}} = -\omega^2\mathbf{x}$.

6. Same as Problem 5 for the equation, in $\mathcal{R}^2$, $\dot{x} = ax + y, \dot{y} = -x + ay$, discussing the result in terms of $a \in \mathcal{R}$.

7. If Eq. (2.6.1) is normal and reversible, prove that it is also normal in the past. Show that the flow $(S_t)_{t\geq 0}$, solving it for $t > 0$, can be extended to a flow $(S_t)_{t\in\mathcal{R}}$, solving Eq. (2.6.1) for all $t \in \mathcal{R}$: one can define $S_{-t} = S_t^{-1}$, $\forall\, t \geq 0$. In this case, the family $(S_t)_{t\in\mathcal{R}}$ forms a group of maps of $\mathcal{R}^{ds}$ onto itself.

2.7 One-Dimensional Conservative Periodic and Aperiodic Motions

Having completed a general survey on existence, uniqueness, and regularity properties of ordinary differential equations, let us go back to the qualitative theory of motions developing under the action of a purely positional force f considered in §2.1 [see Eq. (2.1.1)]. For such motions the energy conservation theorem was derived (so that they are called "conservative motions" generated by "conservative forces"). The analysis will now concern another qualitative property and study under which circumstances the motions are periodic or aperiodic.

Let V be the potential energy generating the force f (i.e., $f(\xi) = -\frac{dV}{d\xi}(\xi)$ and, in order to have motions defined for all times ("globally defined motions"), suppose that V is bounded below (see Proposition 6). Let $(\eta_0, \xi_0) \in \mathcal{R}^2$, $t_0 \in \mathcal{R}$, and let $t \to \xi(t)$, $t \geq t_0$, be the solution to Eq. (2.1.1) with data:

$$\dot{x}(t_0) = \eta_0, \quad x(t_0) = \xi_0. \tag{2.7.1}$$

If $E = \frac{m}{2}\eta_0^2 + V(\xi_0)$, we can represent graphically the initial datum and the potential as in Fig.2.1.

If ξ_0, as in the picture, is between two contiguous and distinct solutions $x_-(E) < x_+(E)$ of $V(\xi) = E$ and if $\frac{dV}{d\xi}(x_-(E)) < 0$, $\frac{dV}{d\xi}(x_+(E)) > 0$, then by energy conservation the motion $t \to x(t)$ will never leave the interval $[x_-(E), x_+(E)]$. In fact V would be strictly larger than E to the left of $x_-(E)$ and to the right of $x_+(E)$ and, therefore, the motion with energy E would have to have negative kinetic energy when occupying such a position.

Such a trapped motion will be periodic if and only if it takes a finite time for it to run from $x_-(E)$ to $x_+(E)$. This amount of time can be estimated easily by the quadrature formula (2.1.8), p. 12.

If $x_-(E)$ or $x_+(E)$ or both do not exist, the above argument says that the motion may be unbounded. The above argument also does not give any

precise predictions when the derivative of V vanishes in at least one of the two points $x_-(E), x_+(E)$.

The following proposition provides a general result and in its proof all the above problems are implicitly or explicitly solved.

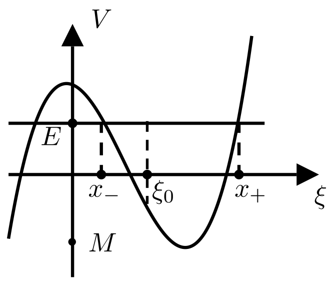


Figure 2.1.: Two contiguous roots of $V(x) = E$.

11 Proposition. *The motion $t \to x(t)$, $t \in [t_0, +\infty)$ of $m\,\ddot{x} = -\frac{dV}{dx}(x)$ with initial datum (2.7.1) is periodic with a positive minimal period if and only if ξ_0 is between two adjacent roots $x_- < x_+$ of $V(\xi) = E$, where the derivative of V is, respectively, negative and positive.*

PROOF. Suppose that $V(x_\pm) = E$, $-V'(x_-)$ and $V'(x_+) > 0$, and $V(\xi) < E$ for $\xi \in (x_-, x_+)$ and let $\xi_0 \in [x_-, x_+]$. As already noticed it must be that, $\forall\, t \geq t_0$, $x(t) \in [x_-, x_+]$. Suppose, first, that $\eta_0 > 0$ and define $t_+ = \{$supremum of the values $t > t_0$ such that $\dot{x}(\tau) > 0$ for all $\tau \in [t_0, t)\}$. From energy conservation, one deduces:

$$\dot{x}(t) = +\sqrt{\frac{2}{m}(E - V(x(t)))}, \qquad t_0 \leq t < t_+, \tag{2.7.2}$$

where the sign in front of the square root comes from the continuity of $\dot{x}(t)$ and from $\eta_0 > 0$. To estimate t_+, remark that Eq. (2.7.2) implies:

$$t - t_0 = \int_{\xi_0}^{x(t)} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}}, \qquad t_0 \leq t < t_+, \tag{2.7.3}$$

If we show that $\lim_{t \to t_+} x(t) = x_+$, it will follow from Eq. (2.7.3) that

$$t_+ - t_0 = \int_{\xi_0}^{x_+} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}} \tag{2.7.4}$$

which can be used to estimate t_+ and to conclude that $t_+ < +\infty$: i.e., the point reaches x_+ in a finite time.

Once the point reaches x_+, it cannot stay there since $f(x_+) = -\frac{dV}{dx}(x_+) < 0$ and, therefore, $\dot{x}(t_+) < 0$. This means that $\dot{x}(t) < 0$, $x(t) < x_+$ in a right-hand neighborhood of t_+ , by Lagrange's theorem. We can then repeat the already used argument to deduce that:

$$\dot{x}(t) = -\sqrt{\frac{2}{m}(E - V(x(t)))}, \qquad \forall\, t \in [t_+, t_-), \tag{2.7.5}$$

where t_- is analogously defined as $t_- = \{$supremum of the values $t > t_+$ such that $\dot{x}(t) < 0$ for all $\tau \in (t_+, t)\}$. Proceeding as before, we shall show that

$$t_- - t_+ = \int_{x_-}^{x_+} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}} \tag{2.7.6}$$

The same arguments can be again repeated and, therefore, after a suitable time $t' - t_-$, the point will again go through ξ_0 with positive velocity and

$$\int_{x_-}^{\xi_0} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}} \tag{2.7.7}$$

By Proposition 10, p. 35, from now on the motion will identically repeat itself: i.e., $x(t+T) \equiv x(t)$, $\forall\, t \geq t_0$, if T is the sum of the time intervals of Eqs. (2.7.4), (2.7.6), and (2.7.7):

$$T = 2\int_{x_-}^{x_+} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}} \tag{2.7.8}$$

hence, the motion will be periodic and T will be, by construction, its minimal period. It remains to show that $\lim_{t\to t_+} x(t) = x_+$ and that $t_+ < +\infty$.

Since $\dot{x}(t) \geq 0$, $\forall\, t \in [t_0, t_+)$, the limit $\lim_{t\to t_+} x(t) = \overline{x}$ exists and it is approached monotonically. Then, if $\overline{x} < x_+$, it would follow that $\overline{v} = \lim_{t\to t_+} \dot{x}(t) = \left(\frac{2}{m}(E - V(\overline{x}))\right)^{\frac{1}{2}} > 0$; hence, $\dot{x}(t)$ would be > 0 in the right-hand neighborhood of t, if $t_+ < +\infty$, against the very definition of t_+ or, if $t_+ = +\infty$, this would mean that $\overline{x} = +\infty$ against $\overline{x} \leq x_+$. Hence, $\overline{x} = x_+$ and Eq. (2.7.4) holds.

To show that Eq. (2.7.4) also implies $t_+ < +\infty$, apply Lagrange's theorem to infer that there is a point $\widetilde{x} \in (\xi_0, x_+)$ such that for all $\xi \in (\widetilde{x}, x_+)$:

$$E - V(\xi) \geq E - V(x_+) - \frac{1}{2}\frac{dV}{d\xi}(x_+)\,(\xi - x_+) \equiv \frac{f(x_+)\,(\xi - x_+)}{2} \tag{2.7.9}$$

because $E = V(x_+)$ and $f(x_+) = -\frac{dV}{d\xi}(x_+) < 0$ and $(E - V(\xi)) - (E - V(x_+)) - f(x_+)(\xi - x_+)$ is infinitesimal of higher order in $(\xi - x_+)$ as $\xi \to x_+$. Therefore,

$$t_+ - t_0 \leq \int_{\xi_0}^{\widetilde{x}} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}} + \int_{\xi_0}^{\widetilde{x}} \frac{d\xi}{\sqrt{\frac{2}{m}\frac{f(x_+)}{2}(\xi - x_+)}} < +\infty \tag{2.7.10}$$

since the first integral is finite because $\max(2(E - V(\xi))/m)^{-1} < +\infty$ in $[\xi_0, \widetilde{x}]$, while the second integral is also finite (an explicit computation). The

alternatives initially set aside, namely $\eta_0 < 0$ or $\eta_0 = 0$ (i.e., $\xi_0 = x_\pm$) are reduced to the one just treated.

Finally, the cases $f(x_+) = 0$, or $f(x_-) = 0$, or $f(x_+) = f(x_-) = 0$, or x_+ , or x_- not existing have to be discussed and shown to give rise to motions not periodic or with period 0. This last case is realized if (and only if) $\eta_0 = 0$ and $f(\xi_0) = 0$: one says that ξ_0 is an "equilibrium point". Among the remaining cases consider, as an example, the case $\eta_0 > 0$, $f(x_+) = -(dV/d\xi)(x_+) = 0$. Proceeding as before, it is found that t_+ is still given by Eq. (2.7.4). This time, however, to estimate t_+ Eq. (2.7.9) must be improved by using Taylor's formula to second order, since $f(x_+) = 0$. If $f'(x_+) = (df/d\xi)(x_+)$, it is:

$$E - V(\xi) = \frac{1}{2} f'(x_+)(\xi - x_+)^2 + o\big((\xi - x_+)^2\big) \tag{2.7.11}$$

because the left-hand side vanishes together with its first derivative in x_+. Hence $\exists \widetilde{x}' \in [\xi_0, x_+)$ such that, if $f'(x_+) > 0$,

$$E - V(\xi) < f'(x_+)(\widetilde{x} - x_+)^2 + o\big((\xi - x_+)^2\big) \tag{2.7.12}$$

Thus, if $f'(x_+) > 0$, we deduce from Eqs. (2.7.4) and (2.7.2):

$$t_+ - t_0 > \int_{\widetilde{x}'}^{x_+} \Big(\frac{2}{m} \frac{f'(x_+)}{2} (\xi - x_+)^2\Big)^{-\frac{1}{2}} d\xi = +\infty. \tag{2.7.13}$$

The case $f'(x_+) = 0$ is treated likewise, as, in this case, $E - V(\xi)$ is infinitesimal of higher than second order in $\xi - x_+$ and an inequality like Eq. (2.7.12) holds, therefore, with $f'(x_+)$ replaced, say, by 1.

The case $f'(x_+) < 0$ is impossible if $\eta_0 > 0$ (since this would mean that x_+ is a minimum for V). mbe

For future reference let us state the following obvious proposition.

12 Proposition. *If $\xi_0 \in \mathcal{R}$, the constant function $t \to x(t) \equiv \xi_0$ solves to Eq. (2.1.1) if and only if ξ_0 is a stationary point for the potential energy V.*

2.7.1 Exercises and Problems

1. Estimate the period of the motions indicated below with an error rigorously bounded by 60%:

$$\begin{aligned}
x =& x(x-1), && x(0) = 0, && \dot{x}(0) = \frac{1}{\sqrt{6}} \text{ or } \frac{1}{2},\\
\dot{x} =& -2x - 4x^3, && x(0) = 1/\sqrt{2}, && \dot{x}(0) = 0\\
\dot{x} =& -x^3, && x(0) = 0, && \dot{x}(0) = 1,\\
\dot{x} =& -x/(1+x), && x(0) = 0, && \dot{x}(0) = \frac{1}{2},\\
\dot{x} =& \log(1+x), && x(0) = \frac{1}{2}, && \dot{x}(0) = 0.
\end{aligned}$$

(*Hint:* Show first: $I = \int_{x_-}^{x_+} \frac{dx}{\sqrt{(x_+ - x)(x - x_-)Q(x)}} \le I_0 = \frac{1}{Q(\xi)} \int_{x_-}^{x_+} \frac{dx}{\sqrt{(x_+ - x)(x - x_-)}} \equiv \frac{\pi}{Q(\xi)}$, where ξ is any point in $[x_-, x_+]$, with an error $\delta = \frac{|I - I_0|}{I} \le \left[\frac{\max\sqrt{Q}}{\min\sqrt{Q}} - 1\right]$.)

2. Find, if they exist, values of E to which correspond aperiodic motions for the equations in Problem 1, and for: $\ddot{x} = -xe^{-x^2}$; $\quad \ddot{x} = -\sin x$.

3. Same as Problem 1 with an error rigorously bounded by 10% or 1%, using a desk computer.

4. Find whether the motions associated with the second equation in Problem 1 admit a motion with period $T = 10$ and, if it exists, estimate within 20% the amplitude of such a motion.

5. Show that the period of the motion of total energy E verifying $\dot{x} = -x^3$ has a period $T(E)$ proportional to $E^{-\frac{1}{4}}$ if the potential energy is defined as $V(\xi) = \frac{1}{4}\xi^4$. Show that the proportionality constant is $2\int_{-1}^{1}(1-\xi^4)^{-\frac{1}{2}}d\xi$ (*Hint*: Write the formula of quadrature for T, Eq. (2.7.8), and change variable as $\xi \longrightarrow \xi E^{-\frac{1}{4}}$.)

6. Show that the period of the motion with energy E verifying $m\ddot{x} = -(dV/dx)(x)$, with V such that $V(0) = 0, V'(0) = 0, V''(0) > 0$, is such that $\lim_{E\to 0^+} T(E) = 2\pi\left(\frac{m}{V''(0)}\right)^{\frac{1}{2}}$. (*Hint:* see hint to problem 1).

7. Let $\xi \longrightarrow V(\xi)$ be a C^∞ convex even function vanishing at the origin. Let

$$\overline{V}(\xi) = \frac{1}{2}\sigma\,\xi^2 \stackrel{def}{=} \frac{1}{2}\big(\sup_{\xi'} V''(\xi')\big)\,\xi^2$$

Consider a motion, associated with the potential energy V, having total energy E. Show that its period is larger than the period of the motions with potential energy $\overline{V}$.

8. Suppose that $V(\xi) = |\xi|^\alpha$, $\alpha > 1$, and show that the period of the motion with energy E is proportional to $E^{\frac{1}{\alpha}-\frac{1}{2}}$ (see Problem 5).

9. Find the limit as $E \longrightarrow +\infty$ the period of the motion with energy E developing with potential energy $V(\xi) = \frac{1}{2}\xi^2 + \frac{1}{4}\xi^4$

10. Same as Problem 9 with V such that $V(\xi) = V(-\xi)$, $\lim_{\xi\to\infty} \frac{V(\xi)}{\xi^2} = +\infty$.

11. Same as Problem 9 with V such that $V(\xi) = V(-\xi)$, $\lim_{\xi\to\infty} \frac{V(\xi)}{\xi^2} = 0$, $\lim_{\xi\to\infty} V(\xi) = +\infty$.

2.8 Equilibrium: Stability in the Absence of Friction

In the proof of Proposition 12, p. 39, it has been remarked that stationary solutions of $m\,\ddot{x} = f(x)$, i.e., solutions like $t \to \xi_0 =$ constant, correspond to the stationary points of the potential energy function V. In such positions, "equilibrium positions", the exerted force vanishes. It is also possible to further distinguish the equilibrium points on the basis of a qualitative property: the stability of their equilibria. Stability is an empirical notion susceptible to

assuming different precise meanings, depending on the particular problem where it appears necessary to study the stability of an equilibrium point.

It is therefore useful to provide several different definitions of stability for an equilibrium point, leaving to the imagination of the reader the identification of different types of problems for which such types of notions might be relevant. A deeper analysis of the stability notion will be found in Chapter 5, which is entirely devoted to stability theory.

In the following, x_0 shall denote an equilibrium point for $m\ddot{x} = f(x)$ under the assumption that f is generated by a C^∞ potential V bounded from below (so that the equation of motion is normal, see Proposition 6, p.29).

8 Definition. *x_0 is a stable equilibrium position if there is a function $\varepsilon \to a(\varepsilon) \leq +\infty$ defined for $\varepsilon > 0$ and infinitesimal as $\varepsilon \to 0$, such that every motion following an initial condition $x(0) = x_0$, $|\dot{x}(0)| \leq \varepsilon$ has the property:*

$$|x(t) - x_0| < a(\varepsilon), \qquad \forall\, t \geq 0 \tag{2.8.1}$$

Observations.

(1) In other words, x_0 is a stable equilibrium position if a point mass placed in x_0 with small velocity stays indefinitely close to x_0 and the smaller $\dot{x}(0)$, the closer it will stay.
(2) The fact that $a(\varepsilon)$ might be $+\infty$ means that we admit the possibility that initial data whose velocity $\dot{x}(0)$ is too large may originate motions which travel indefinitely far from x_0. Equation (2.8.1) is really a condition which is relevant only for ε small.
(3) The choice of $t_0 = 0$ as initial time is irrelevant since the equation of motion is autonomous.

In most applications it is by no means sufficient to know that x_0 is a stable equilibrium position in the sense of Definition 8. For instance, it is sometimes necessary that the function $a(\varepsilon)$, which could be called the "tolerance" function, has a preassigned structure. This leads to the following definition:

9 Definition. *Given a function of the variable $\varepsilon > 0$, $\varepsilon \to b(\varepsilon) < +\infty$ (not necessarily infinitesimal as $\varepsilon \to 0$), one says that x_0 is a stable equilibrium position "with tolerance b" if the motion $t \to x(t)$, $t \geq 0$, following an initial condition $x(0) = x_0$, $|\dot{x}(0)| \leq \varepsilon$ is such that*

$$|x(t) - x_0| < b(\varepsilon), \qquad \forall\, t \geq 0. \tag{2.8.2}$$

Observations.

(1) Definition 9 differs from Definition 8 because $\varepsilon \to b(\varepsilon)$ is a priori given and also because $b(\varepsilon)$ is not necessarily infinitesimal as $\varepsilon \to 0$.

(2) Obviously one can also give other analogous definitions where the "perturbed" initial data look like $x(0) = x_0 + \varepsilon, \dot{x}(0) = 0$, or some other.

Avoiding formalization of the possibilities hidden in observation 2, some stability criteria will be discussed. A well-known simple criterion for stability in the sense of Definition 8 is stated in the following proposition and it will suggest studying a third stability definition involving the introduction of novel interesting ideas, see §2.9.

13 Proposition. *If x_0 is a strict minimum for the potential energy function V, then x_0 is a stable equilibrium point in the sense of the Definition 8.*

Proof. Let $E_\varepsilon = \frac{1}{2}m\varepsilon^2 + V(x_0)$ be the total energy of the initial datum $x(0) = x_0, \dot{x}(0) = \varepsilon$. By assumption, x_0 is a point of strict minimum for V, see Fig. 2.2, i.e., $V(\xi) > V(x_0)$ if $\xi \neq x_0$ and $|\xi - x_0|$ is small enough; hence it is possible to define the positions $x_{-,\varepsilon}$, and $x_{+,\varepsilon}$, which are the first root of $E - V(\xi) = 0$ to the left or to the right of x_0, respectively. It is also easy to check that the strict minimum assumption also implies that

$$\lim_{\varepsilon \to 0} x_{\pm,\varepsilon} = x_0 \tag{2.8.3}$$

and also that $x_{+,\varepsilon}$ and $x_{-.\varepsilon}$ are, respectively, monotonically increasing and decreasing in ε. For large ε, it might happen that $E_\varepsilon - V(\xi)$ does not have one of the two roots $x, -, \varepsilon$ or $x_{+,\varepsilon}$ or both. In this case define $x_{-,\varepsilon} = -\infty$ or $x_{+,\varepsilon} = +\infty$.

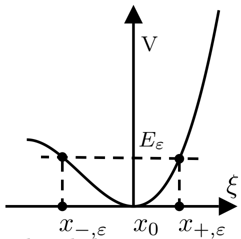


Fig.2.2: A minimum of the potential and the two points $x_{\pm\varepsilon}$.

Then if one sets

$$a(\varepsilon) = \max_{\sigma=\pm 1} |x_{\sigma,\varepsilon} - x_0|, \tag{2.8.4}$$

Eq. (2.8.1) is verified for the motions $t \to x(t)$ such that $x(0) = 0, |\dot{x}(0)| < \varepsilon$, by the arguments of Proposition 11, p. 37, mbe

Observations.

(1) The proof method of Proposition 13 allowed us to define, in fact, the minimal tolerance function; i.e., Eq. (2.8.4). It is therefore easy to provide stability criteria in the sense of Definition 9 by using the preceding proof, under the assumptions of Proposition 13.
(2) Note that if $\frac{d^2V}{d\xi^2}(x_0) > 0$ the function $a(\varepsilon)$ in Eq. (2.8.4) is of $O(\varepsilon)$.

2.8.1 Exercises and Problems

1. Determine the stable equilibrium positions in the sense of Definition 9 with tolerance functions:

$$b(\varepsilon) = \frac{1}{2} + \varepsilon \qquad \text{or} \qquad \begin{cases} b(\varepsilon) = 3\varepsilon & \text{for } \varepsilon < \frac{1}{5}, \\ b(\varepsilon) = +\infty & \text{for } \varepsilon \geq \frac{1}{5}, \end{cases}$$

for a unit mass point acted upon by a force with potential energy

$$V(\xi) = \xi(\xi - 1), \quad \text{or} \quad \log(1 + \xi^2), \quad \text{or} - \sin\xi, \quad \textit{or} \quad \frac{1}{2}\xi^2 e^{-\xi^2}$$

2. Show that not all the stable equilibrium positions for $V(\xi) = (\sin\xi^2)e^{-\xi^2}$ have tolerance $b(\varepsilon) = \frac{1}{2}$ if $\varepsilon \leq 1$ and $b(\varepsilon) = +\infty$ if $\varepsilon > 1$.

3.* Show that the potential energy V defined by

$$V(\xi) = e^{-1/|\xi|}\Big(\xi^2 + (\sin\frac{1}{\xi})^2\Big), \ \xi \neq 0$$

and $V(0) = 0$ has infinitely many stable equilibrium positions in the sense of Definition 8. (*Hint*: Show that $V'(\xi)$ is infinitely many times positive and negative near zero.)

2.9 Stability and Friction

A further alternative definition of an equilibrium point x_0 for a force law $f \in C^\infty(\mathcal{R})$, with potential energy V bounded from below, comes from the remark that, in practice, when x_0 is a stable equilibrium position, then, under a small perturbation of the equilibrium state, the point mass moves away from x_0 to return eventually to x_0 with essentially zero velocity. As it really happens when a pendulum is slightly deflected from its equilibrium position.

To give a mathematically precise meaning to the stability criterion that seems to emerge from these considerations, it is necessary to formulate a precise definition of the term "friction".

An accurate analysis of the friction phenomenon could be found in physics and engineering textbooks: here it will be enough to remark that, empirically, a friction force acts "against the motion"; then one understands why a mathematical model for a friction force is that of a force law depending on the position x and, mainly, on the velocity $\dot{x}$ of the point mass in such a way to have a sign systematically opposite to that of $\dot{x}$.

The simplest model describes the friction force A in terms of a nonnegative C^∞ function $(\eta, \xi) \to \alpha(\eta, \xi)$ on $\mathcal{R}^2$ as:

$$A(\dot{x}, x) = -\dot{x}\,\alpha(\dot{x}, x) \tag{2.9.1}$$

with α verifying the further property that $\alpha(\eta, \xi) \neq 0$ for $\eta \neq 0$; i.e., friction is absent only if the point is standing still. There are, however, phenomena for which this is not a good model, like the so-called "static friction" cases

(which are modeled by discontinuous friction forces). Remarkable examples are: “linear friction”,

$$A(\dot{x}, x) = -\lambda \dot{x}, \qquad \lambda > 0; \tag{2.9.2}$$

“cubic” friction,

$$A(\dot{x}, x) = -\lambda \dot{x}\,(1 + \lambda' \dot{x}^2), \qquad \lambda, \lambda' > 0 \tag{2.9.3}$$

and “quadratic friction”,

$$A(\dot{x}, x) = -\lambda \dot{x}\,(1 + \lambda' \dot{x}^2)^{\frac{1}{2}}, \qquad \lambda, \lambda' > 0 \tag{2.9.4}$$

The following stability notion can then be formulated

10 Definition. *If x_0 is an equilibrium point for $m\ddot{x} = f(x)$, it will be said “strongly stable” if for small enough ε the motions $t \to x(t)$, $t \geq 0$, with initial data $x(0) = x_0, \dot{x}(0) = \varepsilon$ and described by the (normal) equation*

$$m\,\ddot{x} = -\lambda \dot{x} + f(x) \tag{2.9.5}$$

are such that

$$\lim_{t \to +\infty} x(t) = x_0, \qquad \forall\ \lambda > 0 \tag{2.9.6}$$

Observation. In other words, this means that x_0 is strongly stable if, in the presence of an arbitrarily small friction, an initial datum $x(0) = x_0, \dot{x}(0) = \varepsilon$ produces a motion returning asymptotically to x_0, at least if ε is not too large. The following is a stability criterion in the new sense.

14 Proposition. *Let x_0 be an equilibrium point for $m\ddot{x} = -\frac{dV(x)}{dx}$, with $V \in C^\infty(\mathcal{R})$ bounded from below. Suppose that for $\xi - x_0 \neq 0$ and small enough, the derivative $-f'(\xi) = \frac{d^2V(\xi)}{d\xi^2}$ is positive (“strict convexity of V at x_0”); then x_0 is a strongly stable equilibrium point.*

Observations.

(1) The condition on V is verified if, for instance, V has a strict minimum in x_0 and not all its derivatives vanish in x_0.
(2) The function V defined to be 0 for $\xi = 0$ and, for $\xi \neq 0$:

$$V(\xi) = e^{-1/|\xi|}\big(\xi^2 + (\sin \frac{1}{\xi^2})^2\big) \tag{2.9.7}$$

is a potential energy function to which the criterion of Proposition 14 cannot be applied. One can see that, actually, Eq. (2.9.7) provides a counterexample to the thought that might flash that the above strong stability notion is

equivalent to the one of Definition 8. The origin is, in fact, a stable equilibrium position because of Proposition 13, p. 42, but it is not a strongly stable equilibrium point.
(3) The proof of Proposition 14 is a particular case of quite general technique adaptable to the analysis of various stability problems as it will be seen again in Chapter 5.

PROOF. Intuitively it can be expected that, in presence of friction, energy is no longer conserved: it will, indeed, be shown that the energy of the motion $t \to x(t)$, solution to Eq. (2.9.5), defined as $E(t) \stackrel{def}{=} \frac{1}{2}mx(t)^2 + V(x(t)), t \geq 0$, is a non constant function of t, such that $\lim_{t\to\infty} E(t) = E_0 = V(x_0)$. Since $V(\xi) \geq V(x_0) = E_0$ and in the vicinity of x_0 there is just one point, namely x_0, where $V(\xi) = E_0$, it must follow that $\lim_{t\to+\infty} x(t) = x_0$, if ε is small. To study the energy variation, with time, of a motion verifying Eq. (2.9.5), compute its derivative:

$$\frac{d}{dt}E(t) = \frac{d}{dt}\Big(\frac{m\dot{x}(t)2}{2} + V(x(t))\Big) = \dot{x}\,(m\ddot{x} - f(x)) = -\lambda\dot{x}^2 \leq 0 \qquad (2.9.8)$$

which shows that, in presence of linear friction, energy is nonincreasing (and strictly decreasing when the velocity does not vanish). Therefore, the limit

$$E_\infty = \lim_{t\to\infty} E(t) \geq \inf_{\xi\in\mathcal{R}} V(\xi) > -\infty \qquad \text{exists.} \qquad (2.9.9)$$

Since x_0 is, by the assumption on f', a strict minimum point, there are (if ε is small enough) two points $x_{+,\varepsilon}$ and $x_{-,\varepsilon}$, to the right of x_0 and to the left of x_0, respectively, that cannot be bypassed by the motion with Eq. (2.9.5) and initial datum $x(0) = x_0, \dot{x}(0) = \varepsilon$, because $E(t) \leq E(0), \forall\, t \geq 0$. Figure 2.3 eloquently illustrates this, making it unnecessary to expound further details.

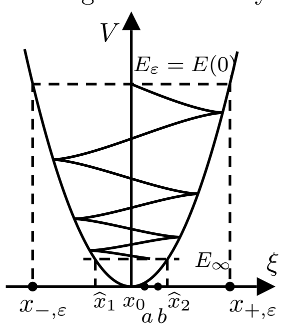


Figure 2.3

Fig.2.3: Decrease in energy as function of time in presence of friction.

Suppose that ε has been chosen so small that, as in Fig.2.3, $f'(\xi) \neq 0$ if $\xi \neq x_0, x_{-,\varepsilon} \leq \xi \leq x_{+,\varepsilon}$: this is possible by the supposed structure of the minimum of V in x_0. Per absurdum, let $E_\infty > E_0$, as in Fig.2.3. Remark first that, as $t \to +\infty$, $\lim_{t\to\infty} x(t) = \widetilde{x}$ must exist. Otherwise, if $\widehat{x}_1 = \liminf_{t\to\infty} x(t) < \widehat{x}_2 = \limsup_{t\to\infty} x(t)$, there would be an interval $[a, b] \subset (\widehat{x}_1, \widehat{x}_2)$, where $\min_{\xi\in[a,b]}(E_\infty - V(\xi)) > 0$, see Fig.2.3. Such an interval would have to be run

infinitely many times as $t \to +\infty$, since $\widehat{x}_1$ and $\widehat{x}_2$ are limit points for $x(t)$; furthermore, when the point mass is in $[a, b]$, its velocity is neither too small nor too large:

$$|x(t)| = \sqrt{\frac{2}{m}(E(t) - V(x(t)))} \geq \sqrt{\frac{2\gamma}{m}} \tag{2.9.10}$$

for some $\gamma > 0$, and

$$|\dot{x}(t)| \leq \sqrt{\frac{2}{m}(E(0) - E_0)}. \tag{2.9.11}$$

Therefore, every time the point mass enters $[a, b]$, it spends therein at least a time T:

$$T = (b - a)\sqrt{\frac{m}{2(E(0) - E_0)}} > 0 \tag{2.9.12}$$

by Eq. (2.9.11) and, therefore [see Eq. (2.9.8)], it loses an amount of energy given, at least, by

$$-\lambda \frac{2}{m} \gamma T \tag{2.9.13}$$

Hence, after infinitely many passages through $[a, b]$, the energy should become $E_\infty = -\infty$, but $E \geq E_0$. Thus the limit $\widehat{x} = \lim_{t\to\infty} x(t)$ exists and $\widehat{x}$ must be one of the two abscissae of the intersections of E_∞, with the graph of V, i.e., in Fig.2.3, one of the two points $\widehat{x}_1$, or $\widehat{x}_2$. Otherwise, $\lim_{t\to\infty} x(t) = \pm(\frac{2}{m}(E_\infty - V(\widehat{x})))^{\frac{1}{2}} \neq 0$ and $x(t)$ could not have a finite limit.[5] This, in turn, implies that $\lim_{t\to+\infty} \dot{x}(t) = 0$.

The last property is, however, in contradiction with the equations of motion (2.9.5) which would imply that

$$\lim_{t\to+\infty} \ddot{x}(t) = \frac{f(\widehat{x})}{m} \neq 0, \tag{2.9.14}$$

i.e., that the limit as $t \to +\infty$ of $\dot{x}(t)$ could not be finite while we proved it to be zero. Hence, E_∞ cannot be larger than E_0, and, then, as already remarked at the beginning of this proof, $\lim_{t\to+\infty} x(t) = x_0$. mbe

2.9.1 Exercises and Problems

1. Show that the equation for the energy variation versus the position is, for the motions verifying $m\ddot{x} + \lambda\dot{x} + V'(x) = 0$, given by:

$$\frac{dE}{dx}(x) = \pm\lambda\sqrt{\frac{2}{m}\Big(E(x) - V(x)\Big)}$$

[5] Exercise: If $\lim_{t\to+\infty} f(t)$ and $\lim_{t\to+\infty} f'(t)$ exist, then $\lim_{t\to+\infty} f'(t) = 0$ (denoting f' the derivative of f).

2.* Consider the motions described by the equations:

$$m\ddot{x} + \lambda\dot{x} + V'(x) = 0, \qquad m\ddot{x} + \lambda\dot{x} + W'(x) = 0,$$

with $\dot{x}(0) = \dot{y}(0) = 0$, $x(0) = y(0) = x_0$ and suppose that for $x_0 \le \xi \le x_1$ one has $0 < -W'(\xi) < -V'(\xi)$.
Denote $v_x(\xi)$ and $v_y(\xi)$ the velocity of the motions x and y, respectively, at their passage through $\xi \in [x_0, x_1]$ and suppose also that it is known that $\dot{x}(t), \dot{y}(t)$ are non-negative for all the times preceding the (respective) time of first passage through x_1.
Show that $v_x(\xi) \ge v_y(\xi)$, $\forall\, \xi \in [x_0, x_1]$. (*Hint*: Use the result of Problem 1 to deduce from $v_x(\xi) = \sqrt{2(E(\xi) - V(\xi))/m}$:

$$\frac{d}{d\xi}(v_x(\xi)^2 - v_y(\xi)^2) = \frac{2}{m}\Big(-\lambda(v_x(\xi) - v_y(\xi)) - V'(\xi) + W'(\xi)\Big).$$

This proves that $(d/d\xi)(v_x(\xi)^2 - v_y(\xi)^2) > 0$ for $\xi > x_0$ and close enough to x_0; hence, for such ξ's, $v_x(\xi) > v_y(\xi)$. If there existed $\overline{\xi} \in (x_0, x_1]$ where $v_x(\overline{\xi}) = v_y(\overline{\xi})$ we could consider the smallest among them: still call it $\overline{\xi}$. Then $(d/d\xi)(v_x(\overline{\xi})^2 - v_y(\overline{\xi}))^2 \le 0$, since $\overline{\xi}$ is the first point where $v_x(\xi) = v_y(\xi)$; but this contradicts the above equation for $v_x(\xi)^2 - v_y(\xi)^2$ since $v_x(\overline{\xi}) = v_y(\overline{\xi})$ while $-V'(\overline{\xi}) + W'(\overline{\xi}) > 0$.)

3.* Consider the case analogous to the one in Problem 2 with initial datum $\dot{x}(0) = \dot{y}(0) = v_0 > 0$.

4.* Formulate and prove results analogous to Problems 2 and 3, when $v_0 < 0, 0 < W'(\xi) < V'(\xi)$.

5. Consider the equation $\ddot{x}+\lambda\dot{x}-f(\xi) = 0$, $x(0) = 0, \dot{x}(0) = 1$ or $\dot{x}(0) = -\sqrt{2/15}$. Determine the limit, as $t \to +\infty$, of $x(t)$ for $\lambda = 50$ and for f with potential energy $V(\xi) = \xi^2(1+\xi)^2$.

6. Same as Problem 5 for $\lambda = 10$ and $x(0) = 0, \dot{x}(0) = 10$.

7. Same as Problem 5 for $V(\xi) = (\xi^2 - 1)(\xi + 2), x(0) = \frac{3}{2}$ and $\dot{x}(0) = 0, \lambda = 4$ or $V(\xi) = \xi^2(\xi + 1)(\xi + 2)$, $\lambda = 1, x(0) = 0, \dot{x}(0) = -\sqrt{2}$.

8. How large should λ be so that the motion verifying $\ddot{x} = -\dot{x} + V'(x)$, with $V(\xi) = \frac{1}{2}\xi^2 e^{-\xi^2}, x(0) = 0, \dot{x}(0) = 10$, is attracted by the origin? (Find a lower bound only.)

9. Show that for λ small enough, the motion in Problem 8 "runs away", i.e., $\lim_{t\to+\infty} x(t) = +\infty$. For such a motion, after an arbitrary choice of λ, estimate the time necessary to reach the point with abscissa $\xi = 10$. (Find an upper and a lower bound.)

2.10 Period and Amplitude: Harmonic Oscillators

In this section a point with mass m is considered subject to a force law f generated by a C^∞ potential energy V such that

$$\begin{aligned} &(i) && V(\xi) = V(-\xi), \\ &(ii) && \frac{dV}{d\xi} \ne 0, \qquad \xi \ne 0 \\ &(iii) && \lim_{\xi\to+\infty} V(\xi) = +\infty \end{aligned} \tag{2.10.1}$$

In §2.7 it was proved that all motions are periodic with period $T < +\infty$.

We now ask whether there exist potential energy functions V verifying Eq. (2.10.1) and generating motions with energy-independent (or amplitude-independent) period.

It is well known that the "elastic energy" $V(\xi) = V(0) + \frac{1}{2}k\,\xi^2$ generates "isochronous" motions of period $T = 2\pi\sqrt{\frac{m}{k}}$, constant as the total energy varies ("harmonic oscillations").

It is remarkable that, in the class (2.10.1)this isochrony is a characteristic property of the harmonic oscillators, ([28]).

15 Proposition. *If all motions developing under the action of a force with potential V verifying Eq. (2.10.1) have the same period, $\exists\, k > 0$ such that*

$$V(\xi) = \frac{1}{2}k\,\xi^2 + V(0). \tag{2.10.2}$$

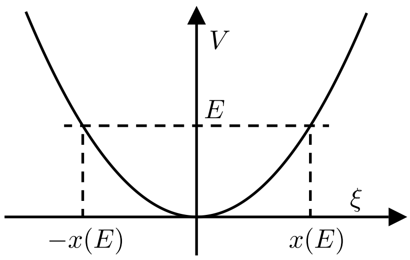


Fig.2.4: A potential satisfying Eq.2.10.1.

Observation. Using the idea involved in the following proof, it is also possible to treat the case when V does not verify (i). See the observations following Corollary 16 below.

PROOF. Let E be the energy of the motion associated with the potential of Eq. (2.10.1) and let $x(E)$ be the corresponding amplitude ($x(E) = x_+$ with the notations of §2.7, see Fig. 2.4). The period of this motion is [see Eq. (2.7.8)]

$$T(E) = 4\int_0^{x(E)} \frac{d\xi}{\sqrt{\frac{2}{m}(E - V(\xi))}}. \tag{2.10.3}$$

Since V is monotonically increasing in ξ for $\xi > 0$, the inverse function to the function V can be defined. Denote it by $v \to \xi(v)$, defined for $v \in [V(0), +\infty)$ and such that $V(\xi(v)) \equiv v$, $\forall\, v \in [V(0), +\infty)$. The second relation in Eq. (2.10.1) implies that $v \to \xi(v)$ is in $C^\infty((V(0), +\infty))$, say, by the implicit function theorem (see Appendix G).

Changing coordinates in Eq. (2.10.4), $\xi \to \xi(v)$, we find:

$$T(E) = 4\int_{V(0)}^{E} \frac{\xi'(v)\, dv}{\left[2(E - V(\xi))/m\right]^{\frac{1}{2}}}, \tag{2.10.4}$$

where $\xi'(v)$ is the derivative of $\xi(v)$ with respect to v. Note that $\xi'(v)$ diverges as $v \to V(0)$, but the divergence is summable in Eq. (2.10.4). Supposing $E \to T(E)$ known for $E \in (V(0), +\infty)$, Eq. (2.10.4) becomes an equation for $\xi(v)$ which can be solved through the following artifice. Multiply Eq. (2.10.4) by $(b-E)^{-\frac{1}{2}}$ and integrate both sides between $V(0)$ and b (assuming that the arbitrary parameter b is chosen larger than $V(0)$):

$$\begin{aligned}\int_{V(0)}^{b} \frac{T(E)\,dE}{\sqrt{b-E}} =&4\sqrt{\frac{m}{2}}\int_{V(0)}^{b} dE \int_{V(0)}^{E} \frac{\xi'(v)\,dv}{\sqrt{(E-v)(b-e)}} \\ =&4\sqrt{\frac{m}{2}}\int_{V(0)}^{b} dv\xi'(v)\Big[\int_{v}^{b} \frac{dE}{\sqrt{(E-v)(b-e)}}\Big].\end{aligned} \tag{2.10.5}$$

The integral in the last parenthesis can be explicitly computed and its value is π ($\forall\ v, b!$). Hence,

$$\int_{V(0)}^{b} \frac{T(E)\,dE}{\sqrt{b-E}} = 4\pi\sqrt{\frac{m}{2}}\big(\xi(b)-\xi(V(0))\big) = 4\pi\sqrt{\frac{m}{2}}\xi(b) \tag{2.10.6}$$

This formula is interesting in itself since it provides the expression of the potential energy "as a function of the period" for all V's verifying Eq. (2.10.1).

When $T(E) = T =$ constant, $\forall\, E \in (V(0), +\infty)$, Eq. (2.10.6) yields

$$\xi(b) = \frac{T}{4\pi}\sqrt{\frac{2}{m}}\,2\,\sqrt{b-V(0)} \tag{2.10.7}$$

which, remembering the definition of $\xi(b)$, means that

$$V(\xi) = \frac{1}{2}m\big(\frac{T}{2\pi}\big)^{-2}\xi^2 + V(0) \tag{2.10.8}$$

mbe

The remark after Eq. (2.10.6) provides the following corollary.

16 Corollary. *Let $E \to T(E), E \in (V(0), +\infty)$, be the period of the motions with energy E developing under the action of a potential verifying Eq. (2.10.1) and let $V(0) = 0$. Then V is given by*

$$\int_{0}^{V(\xi)} \frac{T(E)\,dE}{\sqrt{b-E}} = 4\pi\sqrt{\frac{m}{2}}\,\xi \tag{2.10.9}$$

Observations.

(1) In the above proof, it is necessary that $V(\xi) = V(-\xi)$: if V verifies only (ii) and (iii) of Eq. (2.10.1), then Eq. (2.10.3) is no longer correct and should be replaced by

$$T(E) = 2\int_{x_-(E)}^{x_+(E)} \big(\frac{2}{m}(E - V(\xi))\big)^{-\frac{1}{2}} d\xi, \tag{2.10.10}$$

where $x_+(E), x_-(E)$ are the roots of $E - V(\xi) = 0$ [uniquely defined by Eq. (2.10.1), (ii) and (iii)]. Proceeding as in the proof of Propositions 15 and 16, after splitting Eq. (2.10.10) into two integrals like Eq. (2.10.3) between $x_-(E)$ and 0 and between 0 and $x_+(E)$, it follows:

$$2\pi\sqrt{\frac{2}{m}}(x_+(b) - x_-(b)) = \int_0^b \frac{T(E)\,dE}{\sqrt{b-E}}, \tag{2.10.11}$$

determining $x_+(b) - x_(b)$ in terms of the period function.
(2) Therefore, because of observation (1), there are infinitely many C^∞ functions $\xi \to V(\xi)$ verifying (ii) and (iii) and leading to motions with energy-independent period. They can be visualized by saying that their graphs are obtained by horizontally deforming the parabolae of Eq. (2.10.8), keeping fixed the distances between the values $x_-(E)$ and $x_+(E)$ such that $V(x_\pm(E)) = E$. Hence a necessary and sufficient condition that V, verifying (ii) and (iii) of Eq. (2.10.1), generates isochronous periodic motions is that for all $E > V(0)$,

$$x_+(E) - x_-(E) = k'\sqrt{E - V(0)} \tag{2.10.12}$$

for some $k' > 0$.
(3) Note that Eq. (2.10.9) does not, in general, imply that there is a $V \in C^\infty(\mathcal{R})$ verifying it for arbitrarily given $E \to T(E)$ (see the problems below).

2.10.1 Exercises and Problems

1. Determine the potential V verifying Eq. (2.10.1) and $V(0) = 0$ such that $T(E)$ is $(1+E)$ or $(1+E^2)$ or $\log(1+E)$; check whether $V \in C^\infty(\mathcal{R})$ or $V \in C^\infty(\mathcal{R}/0)$.

2. Let $E \to T(E) > 0$ be a C^∞ function defined for $E > 0$. Suppose that $T(E) = T_0\,(1 + \sum_{k=1}^\infty \tau_k E^k)$ for E small enough and suppose $|\tau_k| \le \varrho^k$ for some $\varrho > 0$. Show that $\xi(b)$ in Eq. (2.10.6) is given by

$$\xi(b) = T_0\sqrt{b}\,\Big(1 + \frac{1}{2}\sum_{k=1}^\infty b^b \tau_k \int_0^1 \frac{x^k\,dx}{(1-x)^{\frac{1}{2}}}\Big)$$

for b small enough.

3. In the context of Problem 2, using the implicit functions theorem (see Appendix G) to invert the function

$$\xi^2 = T_0^2 V\Big(1 + \frac{1}{2}\sum_{k=1}^\infty V^k \tau_k \int_0^1 \frac{x^k\,dx}{(1-x)^{\frac{1}{2}}}\Big)$$

to obtain V as a function of ξ^2 for ξ^2 small, show that there is a $V \in C^\infty(\mathcal{R})$ verifying Eq. (2.10.1) and producing motions with energy E whose period is $T(E)$ for all E small enough; assume $\pi\sqrt{2m} = 1$.

4.* Let $E \to T(E) > 0$ be a C^∞ function defined for $E > 0$. Show that given $N > 0$, there is a C^∞ function A_N such that the function $\xi(b)$ in Eq. (2.10.6) can be expressed as, assuming $\pi\sqrt{2m} = 1$ and E small

$$\xi(b) = T_0\,\sqrt{b}\,\Big(1 + \sum_{k=1}^{N} \widetilde{\tau}_k b^k + b^{N+1} A_{N+1}(b)\Big)$$

where $\widetilde{\tau}_1, \ldots, \widetilde{\tau}_N$ are suitably chosen constants and $T_0 = T(0)$. (*Hint*: Use the Lagrange-Taylor expansion to order N on the left-hand side of Eq. (2.10.6) to express $T(E)$ (see Appendix B).)

5.* Using the result of Problem 4, indicate which, among the following functions $E \to T(E)$, cannot be the function giving the periods of the motions with energy E of some even C^∞ potential: $E \to (1+E)$, $E \to 1 + (\cos E)^2$, $E \to 1 + (\frac{\sin\sqrt{E}}{\sqrt{E}})$, $E \to 1 + \frac{1}{2}(\sin\sqrt{E})^2$, $E \to 1 + \sinh\sqrt{E}$, $E \to 1 + \log(1+E)$, $E \to 1 + \log(1+\sqrt{E})$. (*Hint*: The problem is essentially whether the function (2.10.6) really can be used to obtain b (i.e., V) as a function of ξ which, also, is C^∞.)

6.* Let $V \to \xi(V)$ be defined by

$$\xi(V) = \frac{1}{4\pi}\sqrt{2m}\sqrt{V}\int_0^1 \frac{T(x\,V)\,dx}{\sqrt{1-x}} \qquad V \geq 0$$

obtained from Eq. (2.10.6) by setting $b = V$, $V(0) = 0$, and changing the integration variable. Assume that $E \to T(E)$ is a positive C^∞ function of $E \in [0, +\infty)$. Show that a necessary and sufficient condition for the existence of a potential V verifying Eq. (2.10.1) and producing motions with energy $E \geq 0$ with period $T(E)$ is that $\xi(V) \xrightarrow[V\to+\infty]{} +\infty$, $\xi'(V) > 0$, $\forall\,\xi > 0$. Show also that this happens if $T'(E) \geq 0, \forall\,E > 0$, and does not necessarily happen if one only supposes $T(E)$ bounded for $E > 0$; show, however, that such conditions are only sufficient conditions. (*Hint:* for V near 0 the analysis is in problems 2 through 5 above; if for some V_0 it is $\xi'(V_0) = 0$ the inverse function $\xi \to V(\xi)$ cannot be C^∞ while if $\xi'(V_0) < 0$ it cannot be globally defined for $\xi \in \mathcal{R}$. If $T(E)$ is only supposed bounded a counterexample is $T(E) = 1 + \varepsilon\cos\frac{E}{\varepsilon^2}$ for ε small enough.)

7. Let $V \in C^{(0)}(\mathcal{R})$ verify (i), (ii), and (iii) of Eq. (2.10.1) and suppose that $V \in C^\infty((-\infty, 0) \cup (0, +\infty))$ and $V(0) = 0$. Define $t \to x(t)$, $t \geq 0$ to be a motion generated by V if x is a $C^{(1)}$ function verifying $\dot{x}(t)^2 + V(x(t)) = E > 0$ and $\dot{x}(t)$ changes sign to the right and to the left of any time t when $\dot{x}(t) = 0$. Show that any initial datum $x(0)$, $\dot{x}(0)$ gives rise to a unique motion generated by V and respecting the datum, if $E > 0$.

8.* Show that if in Problem 6 one only drops the condition $T(E) > 0$ replacing it by $T(E) \geq 0$, one has the same results, provided V is allowed to vary in the class of potentials considered in Problem 7.

9. Find a calculation algorithm for the tabulation of a function $\xi \to V(\xi)$ which generates motions with period $\log(1+\sqrt{E})$ with 30% accuracy as E varies in the interval $4 < E < 10$. (*Hint*: Define $T(E)$ "arbitrarily" for $E \not\in [4, 10]$ and use Eq. (2.10.6).)

10. Using a desk computer, actually perform the calculations in Problem 9, drawing (on the screen) the graph of V (without tabulating it) and the graph of the amplitude $x(E), E \in [4, 10]$.

2.11 The Damped oscillator: Euler's Formulae

In §2.7 we saw that the harmonic oscillator is a system with the absolutely remarkable property of exhibiting only periodic motions with the same period.

In this section, and in the following, we shall examine other important properties of harmonic oscillators before dedicating some attention to the study of the stability of such properties with respect to "small" modifications of the force law. Consider a point mass with mass $m > 0$ whose motions are described by the equation

$$m\ddot{x}(t) = -kx(t) - \lambda\dot{x}(t) + \varphi(t), \tag{2.11.1}$$

where $k > 0, a > 0$, and $\varphi \in C^\infty(\mathcal{R})$ is a preassigned function. Equation (2.11.1) is a normal differential equation (see §2.5), as it can be readily verified by multiplying it by $\dot{x}(t)$ and obtaining

$$\frac{d}{dt}E(t) = -\lambda\dot{x}(t)^2 + \dot{x}(t)\varphi(t) \le \max_{\eta\in\mathcal{R}}(-\lambda\eta^2 + \eta\varphi(t)) = \frac{\varphi(t)^2}{4\lambda} \tag{2.11.2}$$

if $E(t) = \frac{1}{2}m\dot{x}(t)^2 + \frac{1}{2}kx(t)^2$. Hence, for all $t > 0$, we find the a priori estimate

$$E(t) = \frac{1}{2}m\dot{x}(t)^2 + \frac{1}{2}kx(t)^2 \le E(0) + \int_0^t \frac{\varphi(t)^2}{4\lambda}\,dt \tag{2.11.3}$$

which implies normality by Proposition 5, p. 28.

Motions described by Eq. (2.11.1) are called "forced oscillations" of a linearly damped harmonic oscillator. In this section we shall study the case $\varphi \equiv 0$, i.e., the equation

$$m\ddot{x} = -\lambda\dot{x} - kx \tag{2.11.4}$$

describing linearly damped oscillators.

The arguments used to prove the strong stability criterion, Proposition 14, p. 44, can be adapted to the particular case of Eq. (2.11.4) and lead to conclude that its motions have a trivial asymptotic behavior as $t \to \infty$: $\lim_{t\to+\infty} x(t) = 0$.

Actually Eq. (2.11.4) can be "explicitly" solved and from the formulae of the solution one gets a very detailed description of the motions as shown by the following proposition.

17 Proposition. *Given $(\eta_0, \xi_0, t_0) \in \mathcal{R}^3$, there exist A_0, A_0' in $\mathcal{R}$ such that the solution of Eq. (2.11.4) with initial datum*

$$\dot{x}(t_0) = \eta_0, \qquad x(t_0) = \xi_0 \tag{2.11.5}$$

can be written as

$$x(t) = e^{-\frac{\lambda}{2m}(t-t_0)}\big(A_0 e^{\frac{\lambda}{2m}\sqrt{1-\frac{4mk}{\lambda^2}}\,(t-t_0)} + A'_0 e^{-\frac{\lambda}{2m}\sqrt{1-\frac{4mk}{\lambda^2}}\,(t-t_0)}\big) \quad (2.11.6)$$

if $\lambda^2 > 4mk$*; or as*

$$\begin{aligned} x(t) = e^{-\frac{\lambda}{2m}(t-t_0)}\big(&A_0 \cos\sqrt{\tfrac{k}{m}(1-\tfrac{\lambda^2}{4mk})}\,(t-t_0) \\ &+ A'_0 \sin\sqrt{\tfrac{k}{m}(1-\tfrac{\lambda^2}{4mk})}\,(t-t_0)\big) \end{aligned} \quad (2.11.7)$$

if $\lambda^2 < 4mk$*; or, if* $\lambda^2 = 4mk$*, as*

$$x(t) = e^{-\frac{\lambda}{2m}(t-t_0)}(A_0 + A'_0\,(t-t_0)) \quad (2.11.8)$$

Observations.

(1) Remark that $\lim_{t\to+\infty} x(t) = 0$ exponentially fast for all solutions.
(2) There are two time scales in the motions described above (they coincide if $\lambda^2 = 4mk$). For small λ (compared with $\sqrt{4mk}$), one time scale is $2m/\lambda$ and the other is $2\pi\sqrt{m/k}$ and $2m/\lambda \gg 2\pi\sqrt{m/k}$. The first time scale controls the damping ("friction time scale") and the other controls the oscillatory motion ("proper time scale") [see Eq. (2.11.7)].
(3) The above solutions can be continued to solutions of Eq. (2.11.4) on the entire time range. However, $\limsup_{t\to-\infty} |x(t)| = +\infty$ unless $x(t) \equiv 0$.

PROOF. A possible proof is by direct verification, i.e., by inserting Eqs. (2.11.6)-(2.11.8) into Eq. (2.11.4) and by checking that in each case the initial data can be satisfied by suitably choosing A_0, A'_0. We present a more instructive proof which illustrates a general method and allows to introduce some new mathematical notions. Look for solutions of Eq. (2.11.4) having the form

$$x(t) = Ae^{\alpha t}, \qquad A \neq 0 \quad (2.11.9)$$

By inserting Eq. (2.11.9) into Eq. (2.11.4), we see that in order that Eq. (2.11.9) be a solution it must be

$$m\alpha^2 + \lambda\alpha + k = 0; \quad (2.11.10)$$

hence, $\alpha = \alpha_+$ or $\alpha = \alpha_-$ with

$$\alpha_\pm = -\frac{\lambda}{2m}\big(1 \pm \sqrt{1 - \frac{4mk}{\lambda^2}}\big) \quad (2.11.11)$$

If $\lambda^2 > 4mk$, there are no problems. For $t \in \mathcal{R}$, setting

$$x(t) = A_0 e^{\alpha_+(t-t_0)} + A'_0 e^{\alpha_-(t-t_0)} \quad (2.11.12)$$

one obtains a solution of Eq. (2.11.4) for all $A_0, A'_0 \in \mathcal{R}$, since Eq. (2.11.4) is a linear homogeneous equation. Imposing the initial conditions yields the system

$$\xi_0 = A_0 + A_0', \qquad \eta_0 = \alpha_+ A_0 + \alpha_- A_0' \tag{2.11.13}$$

whose determinant is $\alpha_+ - \alpha_- = \frac{\lambda}{m}(1 - \frac{4mk}{\lambda^2})^{\frac{1}{2}} \neq 0$. This proves the proposition if $\lambda^2 > 4mk$.

The case $\lambda^2 = 4mk$ can be obtained by first letting $\lambda^2 > 4mk$, solving Eqs. (2.11.4) and (2.11.5), and taking the limit $\lambda^2 \to 4mk$ and using the regularity theorem, Proposition 3, for differential equations.

The determination of A_0 and A_0' from Eq. (2.11.13) gives, for $\lambda^2 > 4mk$,

$$x(t) = \eta_0 \frac{e^{\alpha_+(t-t_0)} - e^{\alpha_-(t-t_0)}}{\alpha_+ - \alpha_-} - \xi_0 \frac{\alpha_- e^{\alpha_+(t-t_0)} - \alpha_+ e^{\alpha_-(t-t_0)}}{\alpha_+ - \alpha_-} \tag{2.11.14}$$

which, as $\lambda^2 \to 4mk$, gives

$$\big(\xi_0 + (\eta_0 + \frac{\lambda}{2m}\xi_0)(t - t_0)\big)\, e^{-\frac{\lambda}{2m}(t-t_0)} \tag{2.11.15}$$

For $\lambda^2 < 4mk$, the roots $\alpha_\pm$ are complex and Eq. (2.11.9) does not directly make sense. However, if we could give a meaning to the exponential of a complex number in such a way that the function $t \to e^{zt}$ has the properties

$$\frac{d}{dt} e^{zt} = z e^{zt}, \qquad \forall\, z \in \mathcal{C} \tag{2.11.16}$$

and, of course, $e^z = \sum_{k=0}^{\infty} z^k/k!$ for z real, we could still take Eq. (2.11.14) as the solution to Eqs. (2.11.4) and (2.11.5). It is natural to define $\forall\, z \in \mathcal{C}$

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!} \tag{2.11.17}$$

since the series is absolutely convergent even if z is complex.

It is then possible to check Eq. (2.11.16) by series differentiation of Eq. (2.11.17) with z replaced by zt: in fact, such a series can be differentiated term by term. Some remarkable properties of e^z are

(i) $$e^z e^{z'} = e^{z+z'}, \qquad e^{\overline{z}} = \overline{e^z} \tag{2.11.18}$$

where the bar denotes complex conjugation. This property can be checked by series multiplication, as for z real, and by conjugation of the series.

(ii) $$e^{x+iy} = e^x\,(\cos y \,+\, i\, \sin y), \qquad \forall\, x, y \in \mathcal{R} \tag{2.11.19}$$

which is checked by recalling the Taylor series for the sine and cosine:

$$e^{x+iy} = e^x e^{iy} = e^x \sum_{k=0}^{\infty} \frac{(iy)^k}{k!} = e^x \sum_{k=0}^{\infty} \Big(\frac{(-1)^k y^{2k}}{(2k)!} + i \frac{(-1)^k y^{2k+1}}{(2k+1)!}\Big) \tag{2.11.20}$$

(iii) By Eq. (2.11.19), one has

$$\cos y = \frac{e^{iy} + e^{-iy}}{2}, \qquad \sin y = \frac{e^{iy} - e^{-iy}}{2i} \tag{2.11.21}$$

Hence, we see that Eq. (2.11.14) gives a solution to Eqs. (2.11.4) and (2.11.5), even if $\lambda^2 < 4mk$, by interpreting the complex exponentials as given by Eqs. (2.11.17) and (2.11.19). Note that Eq. (2.11.14) defines a real function of t, as $\alpha_+ = \overline{\alpha}_-$ and the coefficients of η_0, ξ_0 in Eq. (2.11.14) are therefore real because of the second relation in Eq. (2.11.18). Since, by (2.11.19):

$$\begin{aligned}\mathcal{R}\text{e } e^{\alpha_+ t} &= e^{-\frac{\lambda}{2m}t} \cos \sqrt{\frac{k}{m}(1 - \frac{\lambda^2}{4mk})}\ t,\\ \mathcal{I}\text{m } e^{\alpha_+ t} &= e^{-\frac{\lambda}{2m}t} \sin \sqrt{\frac{k}{m}(1 - \frac{\lambda^2}{4mk})}\ t,\end{aligned} \tag{2.11.22}$$

Eq (2.11.6) follows from Eqs. (2.11.14) and (2.11.22) mbe

Observations

(1) Using the representation (2.11.14) and the complex exponentials, the two cases $\lambda^2 > 4mk$ and $\lambda^2 < 4mk$ are formally unified. This is the first instance, among several that we shall meet, where the use of complex valued functions appears useful and simplifies formulae and calculations even in problems in which one is eventually only interested in "real-valued results".
(2) The formula:

$$e^{x+iy} = e^x\,(\cos y\, +\, i\, \sin y), \qquad \forall\, x, y \in \mathcal{R} \tag{2.11.23}$$

is called "Euler's formula" and it will be widely used in the following.

It is remarkable that the polar representation of a complex number $z = \varrho(\cos\theta + i\sin\theta)$ becomes, because of Eq. (2.11.23):

$$z = \varrho e^{i\theta}, \tag{2.11.24}$$

and also $|e^{iy}| \equiv 1$, $\forall\, y \in \mathcal{R}$ is true and more generally:

$$|e^{x+iy}| = e^x, \qquad \forall\, x, y \in \mathcal{R} \tag{2.11.25}$$

so that $e^z \neq 0$, $\forall\, z \in \mathcal{C}$.

2.11.1 Exercises and Problems

1. Through Euler's formulae prove the "De Moivre formula": i.e., show that for $\forall\, n > 0$ and n an integer, $(\cos\theta + i\sin\sin\theta)^n = (\cos n\theta + i\sin n\theta)$.

2. Through Euler's formulae and the Newton binomial, show that for $n \geq 0$ and integer:

$$(\cos\theta)^n = \Big(\frac{e^{i\theta}+e^{-i\theta}}{2}\Big)^n = \sum_{k=0}^{n}\binom{n}{k}\cos(n-2k)\theta.$$

3. Study the analogue to Problem 2 for $(\sin\theta)^n$.

4. Via Euler's formulae, compute $\sum_{j=0}^{n}\cos j\theta$ using the addition formula for geometric series.

5. Compute

$$\int_0^{2\pi} e^{in\theta}\frac{d\theta}{2\pi}, \quad \int_0^{2\pi}(\cos\theta)^n\frac{d\theta}{2\pi}, \qquad n\in\mathcal{Z}_+.$$

using Euler's formulae and Problem 2.

6. Compute

$$\int_0^{2\pi}(\sin\theta)^n\frac{d\theta}{2\pi}, \quad \int_0^{2\pi}(\sin\theta)^n(\cos\theta)^m\frac{d\theta}{2\pi}, \qquad n,m\in\mathcal{Z}_+.$$

7. Find two linearly independent solutions of $\ddot{x}+\dot{x}+x=0$ and compute their determinant $w(t), t>0$ (see Problem 16 in §2.2).

8. Consider the system of equations in $\mathcal{R}^d$: $\dot{\mathbf{x}} = L\mathbf{x}$, where L is a $d\times d$ matrix $L = (\ell_{ij})_{i,j=1,\ldots,d}$ with constant coefficients. Determine whether there are solutions having the form $x(t) = e^{\alpha t}\mathbf{x}(0)$. Which algebraic equation does α satisfy? Which equation does $\mathbf{x}(0)$ have to verify? (See also Appendices E and F).

9. Apply the method suggested in Problem 8 to find two linearly independent solutions of $\dot{x}=ax+y$, $\dot{y}=-x+ay$ and describe the flow $(S_t)_{t\geq 0}$ in the data space as a varies.

10. Compute the time interval between the n-th and the $(n+1)$-th passage through the origin of the solutions of $\ddot{x}+\dot{x}+x=0$ and $\ddot{x}+\frac{1}{2}\dot{x}+x=0$, in the limit $n\to+\infty$.

2.12 Forced Harmonic Oscillations in Presence of Friction

We now consider Eq. (2.11.1) with $\varphi\neq 0$. Its motions are the "linearly damped harmonic oscillations with forcing term φ".

An obvious but important remark about Eq. (2.11.1) is that its most general solution can be written as the sum of a particular solution $t\to x_{part}(t)$, $t\geq 0$ of Eq. (2.11.1) and of a solution of Eq. (2.11.4), i.e., of the homogeneous equation associated with Eq. (2.11.1). In fact, the linearity of this equation provides that the difference between two of its solutions is a solution of Eq. (2.11.4). Hence, in formulae, a solution $t\to x(t)$, $t>0$, of Eq. (2.11.1) can be written:

$$x(t) = x_{part}(t) + x_0(t), \tag{2.12.1}$$

where $t\to x_0(t)$ is a solution of Eq. (2.11.4).

In 2.11 we saw that $\lim_{t\to+\infty}x_0(t)=0$ and, furthermore, we found explicit expressions for the most general solution $t\to x_0(t)$. Hence, the discussion of

the properties of the motions described by Eq. (2.11.1) is reduced to that of a particular solution of the same equation which we can choose as convenience suggests. This remark is particularly relevant whenever one is interested in the "asymptotic behavior" as $t \to +\infty$, where $t \to x_0(t)$ is infinitesimal.

Let us now describe a method for the construction of a particular solution to Eq. (2.11.1) valid in the interesting though special case when φ is periodic with period $T > 0$.

18 Proposition. *Let $\varphi \in C^\infty(\mathcal{R})$ be a real-valued periodic function with period $T > 0$. Then Eq. (2.11.1) admits a solution with the same period.*

Observation. Consequently, we can say that all the motions described by Eq. (2.11.1) with a periodic forcing term are "asymptotically periodic": this means that there is a periodic solution $t \to x_{per}(t)$, $t \in \mathcal{R}_+$ of Eq. (2.11.1) such that any other solution $t \to x(t)$ has the property $|x(t) - x_{per}(t)| \xrightarrow[t\to+\infty]{} 0$.

Proof. First consider the apparently special cases

$$\varphi(t) = \widehat{\varphi}\cos\frac{2\pi}{T}t \qquad \text{or} \qquad \varphi(t) = \widehat{\varphi}\sin\frac{2\pi}{T}t, \quad \widehat{\varphi} \in \mathcal{R} \tag{2.12.2}$$

and remark that they can be treated simultaneously by solving the equation

$$m\ddot{x} + \lambda\dot{x} + kx = \widehat{\varphi}\, e^{i\frac{2\pi}{T}t} \tag{2.12.3}$$

In fact, the real and imaginary parts of a solution to Eq. (2.12.3) are solutions to Eq. (2.11.1) with φ given, respectively, by the first or the second solution of Eq. (2.11.2) as implied by Euler's formulae $\mathcal{R}\mathrm{e}\ e^{i\frac{2\pi}{T}t} = \cos\frac{2\pi}{T}t$ $\mathcal{I}\mathrm{m}\ e^{i\frac{2\pi}{T}t} = \sin\frac{2\pi}{T}t$.

On the other hand, remembering the properties of the complex exponentials (i.e., $(d/dt)e^{zt} = ze^{zt}$), Eq. (2.12.3) admits a particular periodic solution

$$x_{per}(t) = \frac{\widehat{\varphi}e^{i\frac{2\pi}{T}t}}{-m(\frac{2\pi}{T})^2 + i\lambda\frac{2\pi}{T} + k}. \tag{2.12.4}$$

Hence, the particular cases (2.12.2) are solved by the real and imaginary parts of Eq. (2.12.4), respectively.

To analyze more general cases, linearity of Eq. (2.11.1) can be used again. If this equation is considered with right-hand side $\varphi \in C^\infty(\mathcal{R})$ or $\psi \in C^\infty(\mathcal{R})$ and if $t \to x_\varphi(t)$ and $t \to x_\psi(t)$, $t \in \mathcal{R}_+$, are particular solutions of it, then $t \to x_\varphi(t) + x_\psi(t)$, $t \in \mathcal{R}_+$, is a particular solution of Eq. (2.11.1) with right-hand side $\varphi + \psi$. Consider, then, the case:

$$\varphi(t) = \sum_{n=0}^{N} \widehat{\varphi}_n^{(1)} \cos\frac{2\pi}{T}t + \sum_{n=0}^{N} \widehat{\varphi}_n^{(2)} sin\frac{2\pi}{T}t, \tag{2.12.5}$$

where $\widehat{f}_n^{(1)}, \widehat{f}_n^{(2)}$, $n = 0, 1, 2, \ldots, N$, are real constants. By Euler's formulae, Eq. (2.11.5) can be written as:

$$\varphi(t) = \sum_{n=-N}^{N} \widehat{\varphi}_n e^{i\frac{2\pi}{T}\,n\,t}, \tag{2.12.6}$$

where $\widehat{\varphi}_n$ is defined by

$$\widehat{\varphi}_n = \overline{\widehat{\varphi}_{-n}} = \frac{\widehat{\varphi}_n^{(1)} + i\widehat{\varphi}_n^{(2)}}{2}, \qquad n > 0; \quad \widehat{\varphi}_0 = \widehat{\varphi}_0^{(1)}. \tag{2.12.7}$$

Hence, a particular solution of Eq. (2.11.1) with φ given by Eq. (2.12.5) [or Eq. (2.12.6)] is

$$x_{per}(t) = \sum_{n=-N}^{N} \frac{\widehat{\varphi}_n e^{i\frac{2\pi}{T}\,n\,t}}{-m(n\frac{2\pi}{T})^2 + i\lambda n\frac{2\pi}{T} + k}. \tag{2.12.8}$$

which is real since the addends in Eq. (2.12.8) with index n and $-n$ are complex conjugates because of Eq. (2.12.7). So the proposition is proved when φ is given by Eq. (2.12.5) or Eqs. (2.12.6) and (2.12.7). The same methods can be applied to the case when φ is given by:

$$\varphi(t) = \sum_{n=-\infty}^{\infty} \widehat{\varphi}_n e^{i\frac{2\pi}{T}\,n\,t}, \qquad t \in \mathcal{R}, \qquad \text{with} \tag{2.12.9}$$

$$\widehat{\varphi}_n = \overline{\widehat{\varphi}_{-n}}, \qquad n = 0, 1, \ldots \tag{2.12.10}$$

provided the series (2.12.9) converges well enough so that the function $t \to x_{per}(t)$, $t \in \mathcal{R}$, defined by

$$x_{per}(t) = \sum_{-\infty}^{+\infty} \frac{\widehat{\varphi}_n e^{i\frac{2\pi}{T}\,n\,t}}{-m(n\frac{2\pi}{T})^2 + i\lambda n\frac{2\pi}{T} + k}. \tag{2.12.11}$$

is of class C^∞ and its first and second derivatives (at least) can be computed by summing the corresponding derivatives of the functions in Eq. (2.12.11).

A simple sufficient condition for these properties is that there is a constant c_p such that

$$|\widehat{\varphi}_n| \le \frac{c_p}{(1 + |n|^p)}, \qquad n = 0, \pm 1, \pm 2, \ldots \tag{2.12.12}$$

for all $p > 0$ or, equivalently:

$$\lim_{n\to\infty} |\widehat{\varphi}_n|\,(1 + |n|^p) = 0, \qquad \forall\, p \ge 0 \tag{2.12.13}$$

If Eq. (2.12.12) holds, the series (2.12.9) is uniformly convergent together with the derivative series obtained by differentiating Eq. (2.12.9) term by term an arbitrary number of times. For instance, the series of the k-th derivatives of Eq. (2.12.9) is

$$\sum_{n=-\infty}^{\infty} \widehat{\varphi}_n \big(\frac{2\pi i}{T}\, n\big)^k e^{i\frac{2\pi}{T}\, n\, t}, \tag{2.12.14}$$

and its n-th term has a modulus bounded by

$$|\widehat{\varphi}_n| \big(\frac{2\pi}{T}\, n\big)^k \le \big(\frac{2\pi}{T}\big)^k c_p \, \frac{|n|^k}{(1+|n|)^p} \tag{2.12.15}$$

by Eq. (2.12.12) and by $|e^{i\frac{2\pi}{T}\, n\, t}| \equiv 1$. The right-hand side of Eq. (2.12.15) is t independent and can also be summed over n if one chooses the (arbitrary) parameter $p > k+1$. Then the series differentiation theorems guarantee that Eq. (2.12.9) is a C^∞ function whose derivatives can be computed by "series differentiation".

Hence, the proposition is proved also when φ is given by Eqs. (2.12.9) and (2.12.10) with $\widehat{\varphi}_n$ verifying Eq. (2.12.12), $\forall\, p \ge 0$, i.e., with $\widehat{\varphi}_n$ decreasing faster than any power as $n \to \infty$

The following very important proposition tells us that the last case considered is, actually, the most general and therefore completes our proof.

19 Proposition. *Let $T > 0$ and $\varphi \in C^\infty(\mathcal{R})$ be a periodic function with period T. There exists a unique sequence $(\widehat{\varphi}_n)_{n\in\mathcal{Z}}$ of complex numbers such that*

$$\text{(i)} \qquad \widehat{\varphi}_n = \overline{\widehat{\varphi}_{-n}}, \qquad n = 0, 1, 2, \ldots; \tag{2.12.16}$$

$$\text{(ii)} \qquad \lim_{n\to\infty} (1+|n|)^p |\widehat{\varphi}_n| = 0, \qquad \forall\, p \in \mathcal{Z}_+; \tag{2.12.17}$$

$$\text{(iii)} \qquad \varphi(t) = \sum_{n=-\infty}^{\infty} \widehat{\varphi}_n e^{i\frac{2\pi}{T}\, n\, t}, \qquad \forall\, t \in \mathcal{R}. \tag{2.12.18}$$

The $\widehat{\varphi}_n$ are called the "harmonics" of φ with respect to the period T and

$$\text{(iv)}\widehat{\varphi}_n = \frac{1}{T}\int_0^T \varphi(t) e^{i\frac{2\pi}{T}\, n\, t} dt, \forall\, n \in \mathcal{Z}, \tag{2.12.19}$$

and, finally, $\forall\, s = 0, 1, \ldots$:

$$\frac{d^s\varphi}{dt^s} = \sum_{n=-\infty}^{\infty} \widehat{\varphi}_n \big(\frac{2\pi i}{T}\, n\big)^s e^{i\frac{2\pi}{T}\, n\, t}, \quad \forall\, t \in \mathcal{R} \tag{2.12.20}$$

Observations.

(1) Equation (2.12.17) can also be read as: the sequence $(\widehat{\varphi}_n)_{n\in\mathcal{Z}}$ approaches zero, as $n \to \infty$, "faster than any power". It is equivalent to Eq. (2.12.12).
(2) Proposition 19 implies, via Eqs. (2.12.18) and (2.12.19), that two C^∞ functions periodic with the same period $T > 0$ coincide if and only if all their harmonics relative to the period T coincide.

(3) Proposition 19 is a "structure theorem" on the C^∞-periodic functions on $\mathcal{R}$: it is the "Fourier series theorem".

The proof of this proposition will be given in the next section and it will also conclude the proof of Proposition 18.

2.13 Fourier's series for C^∞-Periodic Functions

Preliminary to the proof of Proposition 19, p. 59, remark that if a function $t \to \varphi(t)$, $t \in \mathcal{R}$, is defined by Eq. (2.12.18) with $(\widehat{\varphi}_n)_{n\in\mathcal{Z}}$ verifying Eqs. (2.12.16) and (2.12.17), then φ is necessarily a C^∞ function, by the series differentiation theorem [see, also, the considerations concerning Eqs. (2.12.14) and (2.12.15)]. Furthermore, since Eq. (2.12.18) is, in this case, uniformly convergent:

$$\int_0^T e^{-i\frac{2\pi}{T}\,n\,t}\varphi(t)\frac{dt}{T} = \sum_{k=-\infty}^{+\infty}\widehat{\varphi}_k\int_0^T e^{-i\frac{2\pi}{T}\,(n-k)\,t}\frac{dt}{T} \tag{2.13.1}$$

by the interchangeability of the integration and the summation operations in uniformly convergent series. However:

$$\int_0^T e^{-i\frac{2\pi}{T}\,(n-k)\,t}\frac{dt}{T} = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases} \tag{2.13.2}$$

as seen by explicit calculation of the integral. Relation (2.13.2) is often written

$$\int_0^T e^{-i\frac{2\pi}{T}\,(n-k)\,t}\frac{dt}{T} = \delta_{nk} \tag{2.13.3}$$

$n, k = 0, 1, \pm 2, \ldots$ with $\delta_{nn} \equiv 1$, $\delta_{nk} \equiv 0$ if $k \neq n$.

Substitution of Eq. (2.13.2) into Eq. (2.13.1) yields

$$\int_0^T e^{-i\frac{2\pi}{T}\,n\,t}\varphi(t)\frac{dt}{T} = \widehat{\varphi}_n, \qquad n \in \mathcal{Z} \tag{2.13.4}$$

which shows that if φ has the form of Eq. (2.12.18) with $(\widehat{\varphi}_n)_{n\in\mathcal{Z}}$ verifying Eq. (2.12.17), then the numbers $\widehat{\varphi}_n$ are uniquely determined by Eq. (2.12.19). If φ is real, then Eq. (2.12.19) [or Eq. (2.13.4)] implies Eq. (2.12.16).

The above considerations show the validity of an "inverse" proposition to Proposition 19 and motivate the validity of Eq. (2.12.19). They are also useful since they allow the introduction of the fundamental relation (2.13.3).

We now give the proof of Proposition 19 §2.12, ("Fourier's theorem").

Proof.. Let $\varphi \in C^\infty(\mathcal{R})$ be a real periodic function with period $T > 0$. Define

$$\widehat{\varphi}_n = \frac{1}{T}\int_0^T e^{-i\frac{2\pi}{T}\,nt}\varphi(t)\,dt, \qquad n \in \mathcal{Z} \tag{2.13.5}$$

It is $\widehat{\varphi}_n = \overline{\widehat{\varphi}_{-n}}$ because T is real. Hence, Eq (2.12.16) holds. To study the asymptotic behavior of $\widehat{\varphi}_n$ as $n \to \infty$, integrate Eq. (2.13.5) by parts.

$$\begin{aligned}\widehat{\varphi}_n =&\frac{1}{T}\big[\frac{e^{-i\frac{2\pi}{T}\,n\,t}}{-i\frac{2\pi}{T}\,n}\varphi(t)\big]_0^T - \frac{1}{T}\int_0^T \frac{e^{-i\frac{2\pi}{T}\,n\,t}}{-i\frac{2\pi}{T}\,n}\varphi'(t)\,dt\\ =&\frac{1}{T}\int_0^T \frac{e^{-i\frac{2\pi}{T}\,n\,t}}{i\frac{2\pi}{T}\,n}\varphi'(t)\,dt\end{aligned} \tag{2.13.6}$$

where φ' denotes the first derivative of φ, and the periodicity of φ has beer used to eliminate the first term in the intermediate relation.

Since φ' is also a T-periodic C^∞ function, and so are the higher derivatives $\varphi'', \varphi''', \ldots, (d^p\varphi/dt^p) \equiv \varphi^{(p)}$, the relation Eq. (2.13.6) can be iterated by again integrating by parts. After p such steps, $p = 0, 1, 2, \ldots$ one finds:

$$\widehat{\varphi}_n = \frac{1}{\big(i\frac{2\pi}{T}\,n\big)^p}\frac{1}{T}\int_0^T e^{-i\frac{2\pi}{T}\,n\,t}\varphi^P(p)(t)\,dt \tag{2.13.7}$$

Hence, if

$$\widetilde{c}_p = \max_{0\le t\le T}|\varphi^{(p)}(t)|, \tag{2.13.8}$$

one has, $\forall\, p = 0, 1, \ldots$:

$$|\widehat{\varphi}_n| \le \big(\frac{T}{2\pi\,|n|}\big)^p\widetilde{c}_p, \qquad \forall\, n \in \mathcal{Z} \tag{2.13.9}$$

which is equivalent to Eq. (2.12.17).

It remains to prove Eq. (2.12.18) with $\widehat{\varphi}_n$, $n \in \mathcal{Z}$, given by Eq. (2.13.5). In fact, the relation (2.12.20) is, as already remarked, a consequence of Eqs. (2.12.18) and (2.12.17). Consider the order N approximation to the series (2.12.18); we elaborate it by using Eq. (2.13.5):

$$\begin{aligned}\sum_{n=-N}^{N} e^{i\frac{2\pi}{T}\,n\,t}\,\widehat{\varphi}_n =& \sum_{n=-N}^{N} e^{i\frac{2\pi}{T}\,n\,t}\int_0^T e^{-i\frac{2\pi}{T}\,n\,\tau}\varphi(\tau)\frac{d\tau}{T}\\ =&\int_0^T \big(\sum_{n=-N}^{N} e^{i\frac{2\pi}{T}\,n\,(t-\tau)}\varphi(\tau)\big)\frac{d\tau}{T}\end{aligned} \tag{2.13.10}$$

which is an identity, $\forall\, \varphi \in C^\infty(\mathcal{R})$.

The summation in the parenthesis in Eq. (2.13.10) is a C^∞ function in t and τ, periodic in both variables with period T, and it has the value $2N+1$ if $\tau = t + mT$, with m an integer. It also has the property

$$\frac{1}{T}\int_0^T \big(\sum_{n=-N}^{N} e^{i\frac{2\pi}{T}\,n\,(t-\tau)}\big)d\tau \equiv 1, \qquad \forall\, t \in \mathcal{R} \tag{2.13.11}$$

which follows from Eq. (2.13.3) by changing $t-\tau$ into t' and by using the mentioned periodicity. Furthermore, the function in parenthesis in Eqs. (2.13.11) and (2.13.10) can be written as

$$1+\sum_{n=1}^{N} e^{i\frac{2\pi}{T}\,n\,(t-\tau)}+\sum_{n=1}^{N} e^{-i\frac{2\pi}{T}\,n\,(t-\tau)} \tag{2.13.12}$$

and the two sums can be "explicitly" summed as geometric sums with ratios $e^{\pm i\frac{2\pi}{T}\,n\,(t-\tau)}$. After a few steps, the result is, for m integer,

$$\sum_{n=-N}^{N} e^{i\frac{2\pi}{T}\,n\,(t-\tau)}=\begin{cases}\frac{\sin(N+\frac{1}{2})\frac{2\pi}{T}(t-\tau)}{\sin\frac{1}{2}\frac{2\pi}{T}(t-\tau)} & \tau\neq t+mT\\ 1+2N & \text{for } \tau=t+mT\end{cases} \tag{2.13.13}$$

Coming back to Eq. (2.13.10) and using Eqs. (2.13.13) and (2.13.11):

$$\begin{aligned}&\sum_{n=-N}^{N}\widehat{f}_n e^{i\frac{2\pi}{T}\,n\,t}=\frac{1}{T}\int_0^T\frac{\sin(N+\frac{1}{2})\frac{2\pi}{T}(t-\tau)}{\sin\frac{1}{2}\frac{2\pi}{T}(t-\tau)}\varphi(\tau)\,d\tau\\ &\equiv\frac{1}{T}\int_0^T\frac{\sin(N+\frac{1}{2})\frac{2\pi}{T}(t-\tau)}{\sin\frac{1}{2}\frac{2\pi}{T}(t-\tau)}(\varphi(t)+\varphi(\tau)-\varphi(t))\,d\tau\\ &=\varphi(t)+\frac{1}{T}\int_0^T\frac{\sin(N+\frac{1}{2})\frac{2\pi}{T}(t-\tau)}{\sin\frac{1}{2}\frac{2\pi}{T}(t-\tau)}(\varphi(\tau)-\varphi(t))\,d\tau\end{aligned} \tag{2.13.14}$$

Hence, to show Eq. (e2.13.8), we have to show that

$$\lim_{N\to\infty}\frac{1}{T}\int_0^T\frac{\sin(N+\frac{1}{2})\frac{2\pi}{T}(t-\tau)}{\sin\frac{1}{2}\frac{2\pi}{T}(t-\tau)}(\varphi(\tau)-\varphi(t))\,d\tau=0 \tag{2.13.15}$$

The reason why this is true is the remark that, at fixed t and $\forall\,m$ integer, the function

$$\tau\to\psi_t(\tau)=\begin{cases}\frac{\varphi(\tau)-\varphi(t)}{\sin\frac{1}{2}\frac{2\pi}{T}(t-\tau)}\cos\frac{1}{2}\frac{2\pi}{T}(t-\tau) & \text{if } t\neq t+mT\\ \frac{T}{\pi}\varphi'(t) & t=t+mT\end{cases} \tag{2.13.16}$$

is just a particular C^∞ function periodic with period t, so that Eq. (2.13.9), and Euler formulae, imply Eq. (2.13.15). The proof of periodicity property of $\psi_t(\tau)$ is left as an exercise (see Problems 1-4 of this section).

By the trigonometric addition formulae, the integral E. (2.13.16) then becomes

$$\frac{1}{T}\int_0^T\Big(\psi_t(\tau)\sin\frac{2\pi}{T}\,N\,(t-\tau)+(\varphi(\tau)-\varphi(t))\,\cos\frac{2\pi}{T}\,N\,(t-\tau)\Big)d\tau \tag{2.13.17}$$

It appears that this expression, via Euler's formulae, is a linear combination of four harmonics of order $\pm N$ of the functions of the τ variable $\tau \to \psi_t(\tau)$ and $\tau \to \varphi(t) - \varphi(\tau)$ which, as discussed above, are C^∞ functions, periodic with period T.

Hence, the integral in Eq. (2.13.17) must tend toward zero faster than any power of N as $N \to \infty$: in fact, the inequalities in Eq. (2.13.9) hold for an arbitrary T-periodic C^∞ function. The same, then, occurs for Eq. (2.13.15), and (2.12.18) is proved. mbe

2.13.1 Exercises and Problems

1. Let $f \in C^\infty(\mathcal{R}), f(0) = 0$. Define $\psi(t) = \frac{f(t)}{t}$, $t \neq 0$, and $\psi(0) = f'(0)$. By applying the Taylor-Lagrange theorem (see Appendix B), show that $\psi \in C^\infty(\mathcal{R})$.

2. In the context of problem 1, show that for $k = 0, 1, \ldots,$

$$\psi^{(k)}(t) = \left(\sum_{h=0}^{k} \frac{(-t)^h}{h!} f^{(h)}(t) \right) \frac{(-1)^k}{t^{k+1} k!}, \qquad \forall\, t \neq 0, \qquad \psi^{(k)}(0) = \frac{f^{(k+1)}(0)}{(k+1)},$$

where the superscript k denotes the kth derivative. (*Hint:* To check that $\psi^{(k)}(t)$ is continuous at $t = 0$ (hence C^∞) remark that the expression in parenthesis is the evaluation of $f(0) = 0$ by Taylor expansion to order k at the point t and evaluated at $-t$; hence vanishes to $O(t^{k+1})$).

3. Show that if $f, g \in C^\infty(\mathcal{R})$ and $g(t_0) = 0, g'(t_0) \neq 0$, the function

$$\psi(t) = \frac{f(t) - f(t_0)}{g(t)}, \quad t \neq t_0, \text{ and } \psi(t_0) = \frac{f'(t_0)}{g'(t_0)}$$

is a C^∞ function in the vicinity of t_0.

4. If $f \in C^\infty(\mathcal{R})$ and is periodic with period T, the function

$$\begin{aligned} \psi_t(\tau) =& \frac{(f(\tau) - f(t)) \cos \frac{\pi}{T}(t - \tau)}{\sin \frac{\pi}{T}(t - \tau)}, \qquad \text{if} \tau \neq t + mT \\ =& \frac{T}{\pi} f'(t), \qquad \text{if} \tau \neq t + mT \end{aligned}$$

if m is any integer, is a C^∞ function of τ and it is periodic with period T. (*Hint*: Use Problem 3.)

5. Using Eq. (2.12.19), compute the Fourier coefficient of order $0, 1, -1$ for the function $f(t) = (1 - \frac{1}{2}\cos t)^{-1}$, thinking of it as a periodic function with period 2π or 4π.

6. Using the Taylor series for the function $(1-\xi)^{-a}$, compute the Fourier series coefficients of the complex-valued functions with period 2π: $f(t) = (1-\frac{1}{2}e^{it})^{-1}$ or $f(t) = (1-\frac{1}{2}e^{it})^{-\frac{m}{n}}$, with $m, n \in \mathcal{Z}$. (*Hint:* $(1-z)^{-a} = \sum_{k=0}^{\infty} \binom{-a}{k}(-z)^k$.)

7. Let, $\forall\, z \in \mathcal{C}$: $\sin z = \frac{e^{iz} - e^{-iz}}{2i}$, $\cos z = \frac{e^{iz} + e^{-iz}}{2}$. Using the Taylor series for the exponential [see Eq. (2.11.17)], determine the Fourier Series coefficients of $f(t) = \sin e^{it}$ or of $f(t) = \cos e^{it}$, $t \in \mathcal{R}$, as 2π-periodic functions or as 4π-periodic functions.

8. Let, $\forall\, z \in \mathcal{C}$, $|z| < 1$: $\log(1+z) = \sum_{k=1}^{\infty} \frac{(-1)^{n+1}}{n} z^n$.. By using the Series expansion for the exponential, Eq. (2.11.17), show that $\exp(\log(l+z)) = 1+z$ and compute the Fourier transform of the 2π-periodic function $f(t) = -\log(1 - \frac{1}{2}e^{it})$.

9. Same as Problem 7 for $f(t) = (1 - \frac{1}{2}\cos t)^{-1}$, $t \in \mathcal{R}$. Estimate $\widehat{f}_2$ up to 10%, i.e., find an expression for $\widehat{f}_n$, but estimate it only for n = 2.

10. Compute the Fourier transform of $f(t) = -\log(1 - \frac{1}{2}\cos t)$ ass a 2π-periodic function. Estimate $\widehat{f}_3$ up to $\%30$.

11. Same as Problem 10 for $f(t) = e^{\cos t}$. Estimate up to 1% the quantities $\widehat{f}_0, \widehat{f}_{\pm 1}$.

12.* Show that a11 the functions in Problems 5-11 have an exponentially decaying Fourier transform. In each case give an estimate of the decay constant.

13. Give an example of a C^∞ function, periodic with period 2π whose Fourier transform does not decay exponentially (*Hint:* First define the transform and then the function, as its sum.)

14. Show that the function $f(t) = \sum_{n=-\infty}^{+\infty} \frac{e^{int}}{(1+n^4)}$ is continuous, term by term differentiable, periodic together with its derivative and with period 2π, but not C^∞.

15.* Analyze critically the proof of §2.13 to deduce that if f is T-periodic continuous and piecewise differentiable with continuous bounded derivatives in each piece, then

$$f(t) = \lim_{N\to+\infty} \sum_{n=-n}^{N} \widehat{f}_n e^{\frac{2\pi}{T} int}$$

with $\widehat{f}_n$ given by $\int_0^T f(t) e^{-\frac{2\pi}{T} int} \frac{dt}{T}$. If f is discontinuous but piecewise continuous with derivatives bounded and continuous in each piece, the preceding formula holds in every continuity point. In the discontinuity points, if $f(t^\pm) \stackrel{def}{=} \lim_{\tau\to t_\pm} f(\tau)$, the series sum is $\frac{f(t_+)+f(t_-)}{2}$ (considering 0 and T as the same point from the point of view of the discontinuities). (*Hint:* To reduce the second part to the first, show the truth of the second part in the case of a function which takes just two values (i.e., which has only two discontinuities being otherwise constant). Then show that any function of the second type is a sum of a function of the first type plus a finite number of piecewise constant functions. Recall that a function f defined on the interval $[a, b]$ is piecewise continuous if a, b can be represented as a union of n closed intervals $[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n]$ and, for every $i = 1, \ldots, n$ the function f coincides in the interior of $[a_i, b_i]$ with a function f_i continuous on the entire interval $[a_i, b_i]$: f may take arbitrary values at the extremes of each interval $[a_i, b_i]$).

2.14 Nonlinear Oscillations. The Pendulum and its Forced Oscillations. Existence of Small Oscillations

In the preceding sections we saw that the asymptotic period of a damped harmonic oscillator is identical to that of the forcing (§2.12). However, the notion of "linear" or "harmonic" oscillator is too rough a notion and, in applications,

a linear oscillator can only appear as a simplified model of some more complex entity.

For instance, very often a linear oscillator appears as a model for the "small oscillations" of a system governed by a nonlinear equation: a prototype of these nonlinear systems is the pendulum.

It is natural to ask the question of the stability of the properties of the solutions to certain classes of equations with respect to the variations of the equations themselves: in fact, it is clear that in applications one shall only "trust" the predictions which do not change by "slightly" changing the models themselves. This is because, as stressed in Chapter 1, there is no "absolutely valid model". As an example of a motion-stability problem in the above sense, we shall now treat some questions concerning the pendulum forced motion; i.e. the motion governed by the (normal) equation:

$$m\ddot{x}(t) + \lambda\dot{x}(t) + k\sin x(t) = f(t), \qquad t \in \mathcal{R}_+ \tag{2.14.1}$$

with λ, m, and $k > 0$ and where $f \in C^\infty(\mathcal{R})$ is a periodic function of period $T > 0$.

In the following, it will be necessary to compare several motions, functions of t, and to fix the ideas we shall adopt, as a measure of magnitude on $[a,b] \subset I$ of a function $\varphi \in C^\infty(I)$ the quantity[6]

$$||\varphi||_{[a,b]} = \sup_{t\in[a,b]} |\varphi(t)| \tag{2.14.2}$$

We now ask if the motions of Eq. (2.14.1) have the following properties

(1) If $t \to x(t), t \in \mathcal{R}_+$, is a motion described by Eq. (2.14.1) and if $x(0), \dot{x}(0), ||f||_{\mathcal{R}_+}$ are small enough, the motion has also oscillations of small amplitude ("existence of small oscillations").

(2) When $||f||_{\mathcal{R}_+}$ is sufficiently small, Eq. (2.14.1) admits a solution with the same period of f.

(3) As $t \to +\infty$, every solution can be asymptotically confused with the periodic solution, in (2) above, provided such a solution exists and the data $x(0), \dot{x}(0)$ are small enough.

In other words, we ask if the above three properties, which have been explicitly or implicitly checked for the forced linear oscillations without restrictions on $\dot{x}(0), x(0), ||f||_{\mathcal{R}_+}$, are still true in a nonlinear case, at least in the small oscillations regime. In this section we analyze problem (1) and introduce the following proposition which "solves" it:

[6] Obviously, there are other possible magnitude measures. Usually the "good" one is determined from the needs of the particular applications. Examples of other measures are

$$\int_a^b |\varphi(t)|dt, \qquad \sup_{t\in[a,b]} (|\varphi(t)| + |\dot{\varphi}(t)|), \qquad \Big(\int_a^b |\varphi(t)|^2 dt\Big)^{\frac{1}{2}}.$$

20 Proposition. *There exist constants* $\gamma, \gamma' > 0$ *such that if* $f \in C^\infty(\mathcal{R})$ *(not necessarily periodic) and* $(x_0, v_0) \in \mathcal{R}^2$*, the motion* $\to x(t)$ *described by Eq. (2.14.1) and following the initial data* $x(0) = x_0, \dot{x}(0) = v_0$*, verifies*

$$||x||_{\mathcal{R}_+} \le \gamma\,(|x_0|+|v_0|+||f||_{\mathcal{R}_+}) \qquad \text{if} \qquad |x_0|+|v_0|+||f||_{\mathcal{R}_+} < \gamma'. \tag{2.14.3}$$

Observations.

(1) Equation (2.14.1) is just one example of a nonlinear equation, chosen among others for its historical and romantic importance. The results and methods that follow apply to much more general equations. The reader will recognize that, in the proof, the key point is that $k \sin \xi - k\xi$; is infinitesimal, as $\xi \to 0$, of higher order in ξ. As an exercise, the reader can, with the obvious modifications, repeat the proof that follows to investigate the validity of the statement identical to Proposition 20 for the equation $m\ddot{x} + \lambda\dot{x} + k\psi(x) = f(t)$, $l > 0$, under the sole assumptions that $\psi \in C^\infty(\mathcal{R}), \psi(0) = 0, \psi'(0) = (d\psi/d\xi)(0) > 0$.
(2) To realize the necessity, in general, of the restriction on $||f||_{\mathcal{R}_+}$ consider the equation

$$m\ddot{+}\lambda\dot{x} + \sin x = \lambda\,\omega + \sin\omega t, \tag{2.14.4}$$

whose solution, among others, $t \to \omega t$ is unbounded. However restrictions on x_0, v_0 are not necessary. In other words, in Proposition 20, one could replace $|x_0| + |v_0| + ||f||_{\mathcal{R}_+} < \gamma'$ with $||f||_{\mathcal{R}_+} < \gamma'$. We have imposed them only for the purpose of simplifying the proof.
(3) The idea behind the proof is to "compare" the solution of Eq. (2.14.1) with the solution of a similar equation where $\sin x$ is replaced by its first-order approximation, namely x. Such comparison will not be "direct", but it will take place by rewriting $k \sin x$ as $kx + k(\sin x - x)$ and considering the function $t \to k(\sin x(t) - x(t))$ as a known function bounded by $k|x(t)|^3/6$, because of the inequality $0 < \xi - \sin\xi \le \xi^3/6$, $\forall\, \xi \in \mathcal{R}_+$.

In this way, one gets a linear equation with forcing term $f(t) - k(\sin x(t) - x(t))$. Solving it "explicitly" (see the following proof), one finds a t-independent relation between the amplitude $M(t) = \max_{0\le\tau\le t} |x(t)| \equiv ||x||_{[0,t]}$ and its cube which, as we shall see, implies that $M(t)$ must stay bounded, $\forall\, t \in \mathcal{R}_+$.

This method of proof is a particular case of a general method to obtain a priori estimates on solutions of nonlinear equations dose to linear ones and, sometimes, it is called the "self-consistency" method. The reader should meditate on the reason for this name after reading the following proof. The self-consistency method will be again used in this book, for instance in the proof of the Lyapunov stability criterion (see §5.4).

Proof. Assume, for simplicity, $\lambda^2 \ne 4mk$. Before analyzing Eq. (2.14.1), it is useful to remark that the equation

$$m\ddot{x} + \lambda\dot{x} + kx = F(t), \qquad t \in \mathcal{R}_+ \tag{2.14.5}$$

with $F \in C^\infty(\mathcal{R})$, admits, among its solutions defined for $t > 0$, the solution

$$p_0(t) = \int_0^t \frac{e^{\alpha_+(t-\tau)} - e^{\alpha_-(t-\tau)}}{\alpha_+ - \alpha_-} F(\tau) \frac{d\tau}{m} \tag{2.14.6}$$

where α_+ and α_- are the two roots of $m\alpha^2 + \lambda\alpha + k = 0$, i.e.,

$$\alpha_\pm \;=\; \tfrac{\lambda}{2m}\left(1 \pm \sqrt{1 - \tfrac{4mk}{\lambda}}\right). \tag{2.14.7}$$

This property can be checked directly by inserting Eq. (2.14.6) into Eq. (2.14.5), and it is a special case of a general property of the linear differential equations which will be illustrated further through exercises and problems at the end of this section.[7]

As already remarked in §2.12, the most general solution to Eq. (2.14.5) will have the form

$$x(t) = \overline{x}(t) + p_0(t), \tag{2.14.8}$$

where $t \to \overline{x}(t)$, $t \in \mathcal{R}_+$, solves Eq. (2.14.5) with $F \equiv 0$. Note also that Eq. (2.14.6) implies that p_0 is real valued, even when $\alpha_\pm$, are complex, provided F is real; also,

$$p_0(0) = 0, \quad \dot{p}_0(0) = 0 \tag{2.14.9}$$

Coming back to Eq. (2.14.1) with initial conditions $x(0) = x_0, \dot{x}(0) = v_0$, we rewrite it as

$$m\ddot{x}(t) + \lambda\dot{x}(t) + k(x(t) \equiv f(t) + k\,(x(t) - \sin x(t))), \tag{2.14.10}$$

and by the preceding remarks, pretending that the right-hand side is a "known function" of t, it is

$$x(t) = \overline{x}(t) + \int_0^t \frac{e^{\alpha_+(t-\tau)} - e^{\alpha_-(t-\tau)}}{\alpha_+ - \alpha_-} \big[f(\tau) + k\,(x(\tau) - \sin x(\tau))\big] \frac{d\tau}{m}, \tag{2.14.11}$$

where $t \to \overline{x}(t)$ is a solution to Eq. (2.14.5) with $F \equiv 0$ and verifying [see Eq. (2.14.9)]

$$\overline{x}(0) = x_0, \qquad \dot{\overline{x}}(0) = v_0. \tag{2.14.12}$$

From §2.12, it follows that ,

[7] Note also that if F is periodic, Eq. (2.14.6) will not be so, in general. Hence, this method for obtaining particular solutions to Eq. (2.14.5) is different from the one in §2.12, valid for periodic F's and based on the Fourier series.

$$\overline{x}(t) = \frac{v_0 - \alpha_- x_0}{\alpha_+ - \alpha_-} e^{\alpha_+ t} - \frac{v_0 - \alpha_+ x_0}{\alpha_+ - \alpha_-} e^{\alpha_- t}, \tag{2.14.13}$$

and since $\mathcal{R}\text{e}\ \alpha_- \le \mathcal{R}\text{e}\ \alpha_+ < 0$, it is $|e^{\alpha_\pm t}| \le e^{\mathcal{R}\text{e}\ \alpha_+ t} \le 1$, $\forall\, t \ge 0$: hence,

$$||\overline{x}||_{\mathcal{R}_+} \le \Big(\frac{|\alpha_+| + |\alpha_-|}{|\alpha_+ - \alpha_-|}|x_0| + \frac{2}{|\alpha_+ - \alpha_-|}|v_0|\Big). \tag{2.14.14}$$

Setting $M(t) = ||x||_{[0,t]}$, we deduce from Eq. (2.14.11), using the inequality

$$0 \le \xi - \sin\xi \le \frac{\xi^3}{6}, \qquad \forall\, \xi \in \mathcal{R}_+, \tag{2.14.15}$$

that, $\forall\, t \ge 0$:

$$|x(t)| \le ||\overline{x}||_{\mathcal{R}_+} + (||f||_{\mathcal{R}_+} + \frac{k}{6}M(t)^3) \int_0^t \frac{d\tau}{m}\, 2\frac{e^{\mathcal{R}\text{e}\ \alpha_+ (t-\tau)}}{|\alpha_+ - \alpha_-|}. \tag{2.14.16}$$

Hence, by integration,

$$|x(t)| \le ||\overline{x}||_{\mathcal{R}_+} + (||f||_{\mathcal{R}_+} + \frac{k}{6}M(t)^3) \frac{2m^{-1}}{|\mathcal{R}\text{e}\ \alpha_+|\,|\alpha_+ - \alpha_-|}. \tag{2.14.17}$$

which implies, by Eq. (2.14.14),

$$|x(t)| \le A + B\, M(t)^3, \qquad t \ge 0 \tag{2.14.18}$$

with

$$A = \Big(\frac{|\alpha_+| + |\alpha_-|}{|\alpha_+ - \alpha_-|}|x_0| + \frac{2}{|\alpha_+ - \alpha_-|}|v_0| + \frac{2m^{-1}||f||_{\mathcal{R}_+}}{|\mathcal{R}\text{e}\ \alpha_+|\,|\alpha_+ - \alpha_-|}\Big), \tag{2.14.19}$$

$$B = \frac{2k}{6m|\mathcal{R}\text{e}\ \alpha_+|\,|\alpha_+ - \alpha_-|}. \tag{2.14.20}$$

It is then immediately seen from Eq. (2.14.18) that the continuity and monotonicity of $M(t) = ||x||_{[0,t]}$ and the arbitrariness of $t \ge 0$ imply

$$M(t) \le A + B\, M(t)^3, \qquad \forall\, t \in \mathcal{R}_+, \tag{2.14.21}$$

and from Eq. (2.14.19), it also follows that

$$M(0) = |x(0)| = |x_0| < A \tag{2.14.22}$$

To complete the proof remark that the graph of the function $M \to A + B\, M^3 - M$ has the form illustrated in Fig. 2.5 if $27BA^2 < 4$. Hence, if $|x_0|, |v_0|, ||f||_{\mathcal{R}_+}$ are small enough so that the latter inequality involving A and B holds [see Eqs. (2.14.19) and (2.14.20)], the equation $A + B\, M^3 - M = 0$ has three real roots $\mu_1(A), \mu_2(A), \mu_3(A)$, with

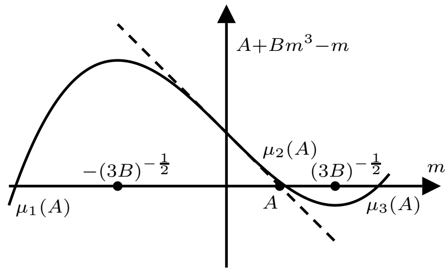


Fig.2.5: Illustration of the bound following from Eq. (2.14.21),(2.14.22).

$\mu_1(A) < 0, 0 < \mu_2(A) < (3B)^{-\frac{1}{2}} < \mu_3(A)$, see Fig.2.5. Furthermore, $A + BM^3 - M > A - M$ for all $M \geq 0$: hence, $\mu_2(A) > A$. Also, if $M \geq 0$, $M < (3B)^{-\frac{1}{2}}$, it follows that $0 \leq A + BM^3 - M \leq A + B(\frac{1}{3B})M - M = A - \frac{2}{3}M$,i.e., $\mu_2(A) \leq \frac{3}{2}A$. So, concluding:

$$A < \mu_2(A) < \frac{3}{2}A \tag{2.14.23}$$

Since the function $t \to M(t)$, $t \in \mathcal{R}_+$, is continuous and verifies Eqs. (2.14.21) and (2.14.22) and $M(t) \geq 0$, it must be

$$M(t) \leq \mu_2(A) \leq \frac{3}{2}A \tag{2.14.24}$$

which concludes the proof. The constant γ' is determined by the condition $27BA^2 < 4$ and γ by Eq. (2.14.24) recalling Eq. (2.14.19). mbe

2.14.1 Exercises and Problems

1. Consider the differential equation $\dot{x} = ax + f(t)$ and show that $p(t) = \int_0^t e^{a\,(t-\tau)} f(\tau) d\tau$ is a solution to it with initial datum $p(0) = 0$, $\forall\, f \in C^\infty(\mathcal{R}), a \in \mathcal{R}$.

2. Let L be a $d \times d$ matrix with constant coefficients and consider the differential equation $\dot{\mathbf{x}} = L\mathbf{x} + \mathbf{f}(t)$, where $\mathbf{f} \in C^\infty(\mathcal{R})$ is an $\mathcal{R}^d$-valued function. Assume that L has d distinct eigenvalues $\lambda_1, \ldots, \lambda_d$ with respective eigenvectors $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \ldots, \mathbf{v}^{(d)}$. Show that if for $\mathbf{w} \in \mathcal{R}^d$, we denote $\alpha_1(\mathbf{w}), \ldots, \alpha_d(\mathbf{w})$ the components of $\mathbf{w}$ on the basis $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$ (see Appendix E), then

$$\mathbf{p}(t) = \int_0^t \sum_{j=1}^d e^{\lambda_j (t-\tau)}\, \alpha_j(\mathbf{f}(\tau))\, \mathbf{v}^{(j)}\, d\tau$$

is a particular solution to the equation, with $\mathbf{p}(0) = \mathbf{0}$. (*Hint:* Note that $\sum_{j=1}^d \alpha_j(\mathbf{f}(\tau))\, \mathbf{v}^{(j)} \equiv \mathbf{f}(t)$ and check the validity of the equation by substitution.)

3.* In the context of Problem 2, Let $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(d)}$ be d linearly independent solutions of the equation $\dot{\mathbf{x}} = L\mathbf{x}$ with initial data $x_j^{(i)}(0) = \delta_{ij}$, $i, j = 1, \ldots, d$. Let $W_{ij}(t) = x_j^{(i)}(t)$, $i, j = 1, \ldots, d$: show that it is the matrix already introduced in Problems 7-9, §2.2 ("wronskian matrix"), verifying $dW/dt = LW$. (Hint: Use the differential equation verified by each row of W.)

4.* In the context of Problems 2 and 3, show that

$$t \to \mathbf{p}(t) = \int_0^t W(t-\tau)\mathbf{f}(\tau)d\tau$$

is a special solution to $\dot{\mathbf{x}} = L\mathbf{x} + \mathbf{f}(t)$ with initial datum $\mathbf{x}(0) = \mathbf{0}$; i.e., it coincides with the one in Problem 2.

5. Apply the method of Problem 2 to find particular solutions to the equation $\dot{x} = -x + y + f_1(t), \dot{y} = -x - y + f_2(t)$.

6. Same as Problem 4 for $m\ddot{x} + \lambda\dot{x} + kx = f(t)$, after reducing it to a first-order system of equations. Consider the case $f(t) = t$. Show that such solution verifies $x(0) = 0, \dot{x}(0) = 0$.

7. Same as Problem 4 for $d^4x/dt^4 - -d^2x/dt^2 + x = t$, after reducing it to a first-order system. Show that such solution verifies $x(0) = 0, \dot{x}(0) = 0, \ddot{x}(0) = 0, x'''(0) = 0$.

2.15 Damped Pendulum: Small Forced Oscillations

We shall now show that the pendulum, as the damped linear oscillator, also admits periodic motions isochronous with the forcing term, at least if the oscillations are small. This solves the problem (2) posed in p. 65, §2.14 Again, the pendulum is selected only for definiteness. The theory developed below is valid for equations obtained from Eq. (2.14.1) by changing $\sin x$ into $\psi(x)$, where ψ is an arbitrary C^∞ function such that $\psi(0) = 0, \psi'(0) > 0$.

Consider the normal equation

$$m\ddot{x} + \lambda\dot{x} + k\sin x = \gamma\, f(t), \qquad t \in \mathcal{R}_+, \tag{2.15.1}$$

$\gamma \in \mathcal{R}, \lambda, m, k > 0, \lambda^2 \neq 4mk$ (for sake of simplicity), then

21 Proposition. *Let $t \to f(t)$, $t \in \mathcal{R}$, be a C^∞ periodic function with period $T > 0$. There exists a periodic motion with period T verifying Eq. (2.15.1), provided γ is small enough.*

Observation. The proof below is based on a very general method used to treat such questions and relying on the implicit functions theorems. Together with Eq. (2.15.1), one considers the "linearized equation"

$$m\ddot{x} + \lambda\dot{x} + k\,x = \gamma f, \tag{2.15.2}$$

which, as shown in §2.12, admits a periodic solution isochronous with f:

$$t \to \overline{x}_\gamma(t) \equiv \gamma\widetilde{x}(t) = \gamma \sum_{n=-\infty}^{+\infty} \frac{\widehat{f}_n e^{\frac{2\pi}{T}int}}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi i}{T}n\lambda + k}, \tag{2.15.3}$$

where $(\widehat{f}_n)_{n\in\mathcal{Z}}$ are the harmonics of f. Then look for a periodic solution to Eq. (2.15.1) having the form

$$t \to x(t) = \gamma\widetilde{x}(t) + y(t), \qquad t \in \mathcal{R} \tag{2.15.4}$$

with initial data:

$$y(0) = \varepsilon, \qquad \dot{y}(0) = \eta \tag{2.15.5}$$

hoping to be able to show that ε, η, and y exist and are infinitesimal as $\gamma \to 0$, of higher order in γ (i.e., hoping that $\gamma\widetilde{x}(t)$ is a very good approximation to $x(t)$ for small γ). The function $t \to x(t)$, solution of Eqs. (2.15.1) and (2.15.5), depends on $\varepsilon, \eta, \gamma$ [in a C^∞ way, by the regularity theorem (see Proposition 3 and Problem 17 of §2.5)]; set

$$x(T) = \gamma\,\widetilde{x}(T) + a(\varepsilon, \eta, \gamma), \qquad \dot{x}(T) = \gamma\,\dot{\widetilde{x}}(T) + b(\varepsilon, \eta.\gamma); \tag{2.15.6}$$

i.e. $y(T) = a(\varepsilon, \eta, \gamma), \dot{y}(T) = b(\varepsilon, \eta, \gamma)$ [see Eq. (2.15.4)]. Therefore, the condition that Eq. (2.15.1) admits a periodic solution with period T can be written (see Proposition 12) as

$$a(\varepsilon, \eta, \gamma) = \varepsilon, \qquad b(\varepsilon, \eta, \gamma) = \eta, \tag{2.15.7}$$

since $\widetilde{x}(0) = \widetilde{x}(T), \dot{\widetilde{x}}(0) = \dot{\widetilde{x}}(T)$ by the periodicity of $\widetilde{x}$.
So the problem of proving Proposition 21 is equivalent to proving the solubility of the implicit functions problem of expressing, from Eq. (2.15.7), ε and η as functions of γ for γ small.

Proof.. Note that the functions $f_1(\varepsilon, \eta, \gamma) = a(\varepsilon, \eta, \gamma) - \varepsilon$ and $f_2(\varepsilon, \eta, \gamma) = b(\varepsilon, \eta, \gamma) - \eta$ are C^∞ functions. To study Eq. (2.15.7), write the equation verified by $t \to y(t)$ defined in Eq. (2.15.4):

$$\begin{aligned} &m\ddot{y}(t) + \lambda\dot{y}(t) + k\,y(t) = k(\gamma\widetilde{x}(t) + y(t) - \sin(\gamma\widetilde{x}(t) + y(t))), \\ &y(0) = \varepsilon, \qquad \dot{y}(0) = \eta \end{aligned} \tag{2.15.8}$$

This equation and the uniqueness theorem for differential equations show that if $\varepsilon = 0, \eta = 0, \gamma = 0$, it follows that $y(t) \equiv 0$, $t \in \mathcal{R}_+$ and, therefore,

$$f_1(0,0,0) = 0, \quad f_2(0,0,0) = 0 \tag{2.15.9}$$

It is then natural to look for solutions of Eq. (2.15.7) near $\gamma = 0$ through the implicit functions theorem (see Appendix G). The solubility condition of Eq. (2.15.7) for small γ is that the Jacobian matrix

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial \varepsilon}(0,0,0) & \frac{\partial f_1}{\partial \eta}(0,0,0) \\ \frac{\partial f_2}{\partial \varepsilon}(0,0,0) & \frac{\partial f_2}{\partial \eta}(0,0,0) \end{pmatrix} \tag{2.15.10}$$

has non vanishing determinant. To compute the derivatives in Eq. (2.15.10), recall that

$$a(\varepsilon,\eta,\gamma) = y(T), \qquad b(\varepsilon,\eta,\gamma) = \dot{y}(T), \tag{2.15.11}$$

where $t \to y(t)$ solves Eq. (2.15.8). Pretending that the right-hand side of Eq. (2.15.8) is a known function of $t \in \mathcal{R}_+$, write

$$\begin{aligned} y(t) =&\overline{y}(t) + \int_0^t \frac{e^{\alpha_+(t-\tau)} - e^{\alpha_-(t-\tau)}}{\alpha_+ - \alpha_-} \cdot \\ &\cdot k\left(\gamma\widetilde{x}(t) + y(\tau) - \sin(\gamma\widetilde{x}(\tau) + y(\tau))\right)\frac{d\tau}{m}, \end{aligned} \tag{2.15.12}$$

along the same lines as the proof in the preceding section, where $t \to \widetilde{y}(t)$ is a solution to

$$m\ddot{\overline{y}} + \lambda\dot{\overline{y}} + k\,\overline{y} = 0, \qquad \overline{y}(0) = \varepsilon,\, \dot{\overline{y}}(0) = \eta, \tag{2.15.13}$$

[see Eq. (2.14.13)]:

$$\overline{y}(t) = \frac{\eta - \alpha_-\varepsilon}{\alpha_+ - \alpha_-}e^{\alpha_+ t} + \frac{\alpha_+\varepsilon - \eta}{\alpha_+ - \alpha_-}e^{\alpha_- t}. \tag{2.15.14}$$

Hence,

$$a(\varepsilon,\eta,\gamma) = \eta\frac{e^{\alpha_+ T} - e^{\alpha_- t}}{\alpha_+ - \alpha_-} + \varepsilon\frac{\alpha_+ e^{\alpha_- T} - \alpha_- e^{\alpha_+ t}}{\alpha_+ - \alpha_-} \tag{2.15.15}$$

$$+ \int_0^T \frac{e^{\alpha_+(t-\tau)} - e^{\alpha_-(t-\tau)}}{\alpha_+ - \alpha_-}\, k\left(\gamma\widetilde{x}(t) + y(\tau) - \sin(\gamma\widetilde{x}(\tau) + y(\tau))\right)\frac{d\tau}{m},$$

and a similar expression can be found for b by differentiating Eq. (2.15.12) with respect to t and setting $t = T$. From Eq. (2.15.15), we can compute the partial derivatives of a with respect to $\varepsilon, \eta, \gamma$ in $(0,0,0)$, without really knowing $y(t)$ (remarkably enough). For instance:

$$\begin{aligned} \frac{\partial a}{\partial\varepsilon}(0,0,0) = &\frac{\alpha_+ e^{\alpha_- T} - \alpha_- e^{\alpha_+ T}}{\alpha_+ - \alpha_-} + \int_0^T \frac{d\tau}{m}\left\{\frac{\alpha_+ e^{\alpha_-(T-\tau)} - \alpha_- e^{\alpha_+(T-\tau)}}{\alpha_+ - \alpha_-}\cdot\right. \\ &\left.\cdot k(1 - \cos y(\tau))\,\frac{\partial y}{\partial\varepsilon}(\tau)\right\} \equiv \frac{\alpha_+ e^{\alpha_- T} - \alpha_- e^{\alpha_+ T}}{\alpha_+ - \alpha_-} \end{aligned} \tag{2.15.16}$$

where $\tau \to y(\tau)$, $\tau \geq 0$, is the solution to Eq. (2.15.8) with $\varepsilon = \eta = \gamma = 0$. Note that $\frac{\partial y}{\partial\varepsilon}(\tau)$ is unknown but is multiplied by zero and, therefore, it is not necessary to know it. Similarly:

$$\frac{\partial a}{\partial\eta}(0,0,0) = \frac{e^{\alpha_+ T} - e^{\alpha_- T}}{\alpha_+ - \alpha_-}, \tag{2.15.17}$$

$$\frac{\partial b}{\partial \varepsilon}(0,0,0) = \alpha_+\alpha_- \frac{e^{\alpha_- T} - e^{\alpha_+ T}}{\alpha_+ - \alpha_-}, \qquad \frac{\partial b}{\partial \eta}(0,0,0) = \frac{\alpha_+ e^{\alpha + - T} - \alpha_- e^{\alpha_- T}}{\alpha_+ - \alpha_-},$$

hence, it is possible to write the matrix J and, with some patience, the algebraic calculations lead to

$$\det J = (\alpha_+ e^{\alpha_+ T} - 1)(e^{\alpha_- T} - 1) \neq 0. \tag{2.15.18}$$

This completes the proof since the implicit functions theorem (Appendix G) implies that Eq. (2.15.7) can be uniquely solved for small γ with $\varepsilon(\gamma), \eta(\gamma)$ of the order $O(\gamma)$.

Actually, the implicit functions theorem implies that the derivatives of $\varepsilon(\gamma), \eta(\gamma)$, with respect to γ at $\gamma = 0$ are proportional to the derivatives of f_1, and f_2 with respect to γ in $\varepsilon = \eta = \gamma = 0$. Since such derivatives can be computed in the same way as those in Eqs. (2.15.16) and (2.15.17) and they turn out to be zero, it also follows that $\varepsilon(\gamma), \eta(\gamma)$ are of the order $O(\gamma^2)$ as expected. mbe

2.15.1 Problems

1. Show that the oscillator $\ddot{x}+\dot{x}+x+x^3 = f(t)$, $f \in C^\infty$, has small oscillations in the sense of Proposition 20. Show that if f has the form $f(t) = \gamma\,\varphi(t)$, $g \in \mathcal{R}$, and φ periodic with period $T > 0$, then for γ small enough the equation admits a periodic solution with period T. (*Hint.*: Go through the proof of Proposition 21, replacing $\sin x$ by $x + x^3$ everywhere.)

2. Show that the motion $\ddot{x} + \dot{x} + x^3 = 0$, $x(0) = 1, \dot{x}(0) = 0$ never goes through the origin as $t \to +\infty$. How does this result depend on the datum? (*Hint:* From $\ddot{x} + \dot{x} + x = x(1 - x^2)$ write $x(t)$ as in Eq. (2.15.12) which will imply $x(t) \geq 0$. To study other initial data x_0 use $\frac{dE}{dx} = -\sqrt{2(E(x) - \frac{x^4}{4}})$, see problem 1, §2.9, and supposing that the first passage time is $t_0 = +\infty$ deduce that this implies $E(0) = \int_0^{x_0} \sqrt{2(E(x) - \frac{x^4}{4}})dx \leq \int_0^{x_0} \sqrt{2(E(0) - \frac{x^4}{4}})dx \leq \sqrt{2E(0)}\,x_0$, i.e. $x_0 \leq 2\sqrt{2}$ and infer that if $x_0 > 2\sqrt{2}$ the point passes through the origin).

3. Same as Problem 2 for $\ddot{x}+3\dot{x}+x+x^3 = O$, $x(0) = \frac{1}{2}, \dot{x}(0) = 0$. (*Hint:* Write the equation as $\ddot{x} + 3\dot{x} + 2x = x(1 - x^2)$ and follow the hint to problem 2).

4.* Consider the oscillator $\ddot{x} + \dot{x} + x^3 = 0$ and find the limit T_∞, as $t \to +\infty$, of the time $T(t)$ elapsing between the two consecutive passages through the origin with positive speed taking place after t. Show that it does not depend on the initial datum (*Answer:* $T = 4\pi/\sqrt{3}$).(*Hint:* Let $x_1, x_2, \ldots$ be the successive maxima of the motion and let $t_1, t_2, \ldots$ be the corresponding times; call $\bar{t}_1, \bar{t}_2, \ldots$ the first passage times through the origin following $t_1, t_2, \ldots$, respectively; then use Eq. (2.15.12) in the intervals $[t_1, \bar{t}_1]$, $[t_2, \bar{t}_2]$ and the fact that $x_i \xrightarrow[i\to\infty]{} 0$.)

5.* Same as Problem 4 for $\ddot{x} + \dot{x} + \tanh x = 0$; discover why T_∞ is the same as that in Problem 4.

6. Examine critically the proof of Proposition 21 to see under which assumptions its conclusions remain valid when $\lambda = 0$. (*Answer:* If and only if $\det J \neq 0$, i.e., if and only if the forcing period T is not an integer multiple of the "proper period $T_0 = 2\pi\sqrt{m/k}$.)

2.16 Small Damping: Resonances

We shall not study the problem (3), p.65, in detail since, in the next few sections, we shall conclude our analysis of the damped motions (in one dimension) by an application where a similar, but more difficult, problem is analyzed. Let us simply formulate a result about problem (3), without proof:

22 Proposition. *Let $f \in C^\infty(\mathcal{R})$ be a periodic function with period $T > 0$ and consider the forced pendulum of Eq. (2.15.1). If γ is small enough, the equation admits one periodic solution $t \to x_p(t)$, $t \in \mathcal{R}_+$, with period T, and every other solution $t \to x(t)$, $t \in \mathcal{R}_+$, with initial datum (x_0, v_0) with $|x_0| + |v_0|$ small enough approaches, exponentially fast, the periodic solution: i.e., there are $C > 0, \mu > 0$ such that*

$$|x(t) - x_p(t)| \leq C\, e^{-\mu t}, \qquad \forall\, t \in \mathcal{R}_+. \tag{2.16.1}$$

Observations.

(1) The proof of this proposition is very similar to that of Proposition 25 on the theory of the clock. The reader will reconstruct it from that proof.
(2) Hence, the small oscillations of nonlinear damped oscillators are qualitatively very similar to those of damped linear oscillators, at least if one is only concerned with properties (1), (2), and (3) selected for discussion at §2.14.

So far, the presence of friction has revealed itself to be essential to the theory (see, however, Problem 6, §2.15). In fact something "goes wrong" as $\lambda \to 0$. This can be seen for the linear oscillators, as it will be briefly discussed in the following. This time, however, consideration of only harmonic oscillators will not just be "for simplicity", but because only in this case will it be possible to obtain something without excessive conceptual and technical difficulties.

In the nonlinear case, the discussion is, surprisingly at first sight, much more involved (and interesting) and, also, the results are unfortunately less detailed and complete than desirable for applications. Some basic ideas and technical tools will be developed in §5.9-§5.12 of Chapter 5.

Actually, contrary to what is sometimes believed, the motion of mechanical systems is much simpler and stable when friction is present than when it is absent. When friction vanishes, the motion becomes very sensitive to the details of the equations of motion and to the initial data, as far as asymptotic behavior is concerned, in this way introducing new difficulties and peculiarly new phenomena. Also, from the mathematical point of view, the frictionless motion theory appears to be deep and rich with connections to the most diverse fundamental problems in analysis and geometry:[8] from number theory to topology to probability theory.

[8] However, at a deeper level of understanding, similar statements could also be made for dissipative systems: a glimpse of how complex they may become is given in §5.8.

Our discussion of the small friction case will be based on the following two linear (normal) equations:

$$m\ddot{x} + \lambda\dot{x} + kx = f \tag{2.16.2}$$

with $\lambda > 0, k > 0, m > 0$ and

$$m\,\ddot{x} + k\,x \;=\; f, \tag{2.16.3}$$

where f is a C^∞ function periodic with period $T > 0$. The discussion will be restricted to the following simple proposition (for the time being).

23 Proposition. *Given $x_0, v_0 \in \mathcal{R}$, let $t \to x_\lambda(t)$, $t \in \mathcal{R}_+$ be the solution to Eq. (2.16.2) with initial data $x_\lambda(0) = x_0, \dot{x}_\lambda(0) = v_0$. Let $t \to x_0(t)$, $t \in \mathcal{R}_+$, be the solution to Eq. (2.16.3) with data $x(0) = x_0, \dot{x}(0) = v_0$, the following results hold:*

$$(i) \qquad \lim_{\lambda\to 0} x_\lambda(t) = x_0(t), \qquad \forall\, t \in \mathcal{R}_+. \tag{2.16.4}$$

(ii) The preceding limit is "uniform as $\lambda t \to 0$": i.e., given $\varepsilon > 0$, there exist $\delta_\varepsilon > 0$, $\lambda_\varepsilon > 0$ such that

$$|x_\lambda(t) - x_0(t)| < \varepsilon \qquad \forall\, \lambda < \lambda_\varepsilon,\ \forall\, t < \delta_\varepsilon \lambda^{-1} \tag{2.16.5}$$

(iii) If T is not an integer multiple of the "proper period" $T_0 = 2\pi\sqrt{m/k}$ of the undamped free harmonic oscillator, one has

$$x_0(t) = A_0 \cos(\frac{2\pi}{T_0} t + \varphi_0) + \sum_{n=-\infty}^{\infty} \frac{\widehat{f}_n e^{\frac{2}{\pi}} T}{-m\big(\frac{2\pi}{T}\big)^2 n^2 + k}, \tag{2.16.6}$$

where A_0, φ_0 are suitable constants and $(\widehat{f}_n)_{n\in\mathcal{Z}}$ are the harmonics of f on the period T: this is the "non resonant case".
(iv) If $T = \overline{n} T_0$ for some integer $\overline{n}$:

$$\begin{aligned} x_0(t) =& A_0 \cos(\frac{2\pi}{T_0} t + \varphi_0) + \sum_{\substack{n=-\infty \\ n\neq\pm\overline{n}}}^{+\infty} \frac{\widehat{f}_n e^{\frac{2\pi i}{T} nt}}{-m\big(\frac{2\pi}{T}\big)^2 n^2 + k} \\ & + 2t\mathcal{R}\mathrm{e}\, \frac{\widehat{f}_n e^{\frac{2\pi i}{T} nt}}{2i\,(\frac{2\pi}{T_0}) m} \end{aligned} \tag{2.16.7}$$

This is the "resonant case".

Observations.

(1) (ii) is particularly significant and says that the smaller the friction, the longer the time during which the friction-driven motion coincides, within a given approximation ε, with the frictionless motion (this time being at least $\delta_\varepsilon/\lambda$). Hence, (ii) strengthens and implies (i).

(2) The above proposition also illustrates the "resonance phenomenon". By what has been seen in §2.12, the solution to Eq. (2.16.2) of interest to us is

$$x_\lambda(t) = A_+ e^{\alpha_+ t} + A_- e^{\alpha_- t} + \sum_{n=-\infty}^{+\infty} \frac{\widetilde{f}_n e^{\frac{2\pi}{T} int}}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi i}{T} n\lambda + k}, \tag{2.16.8}$$

where $v_0 = \dot{x}_\lambda(0), x_0 = x_\lambda(0)$ will determine the constants A_+, A_- and

$$\alpha_\pm = -\frac{\lambda}{2m}\Big(1 \pm i\sqrt{\frac{4mk}{\lambda^2} - 1}\Big) \tag{2.16.9}$$

From Eq. (2.16.8), it immediately follows that as $t \to +\infty$, the asymptotic motion is T periodic and it is given by

$$\overline{x}_\lambda(t) = \sum_{n=-\infty}^{+\infty} \frac{\widetilde{f}_n e^{\frac{2\pi}{T} int}}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi i}{T} n\lambda + k}, \tag{2.16.10}$$

provided the first two "transient terms" in Eq. (2.16.8) are very small: i.e., provided $\lambda t/2m \gg 1$.

If there is $\overline{n}$ such that $T = \overline{n} T_0$, select the two terms in Eq. (2.16.10) with $n = \pm\overline{n}$ and rewrite them as

$$\overline{x}_\lambda = 2\mathcal{R}\mathrm{e}\, \frac{\widehat{f}_n e^{\frac{2\pi i}{T} nt}}{2i\,(\frac{2\pi}{T_0})m} + \sum_{|n|\neq\overline{n}} \frac{\widetilde{f}_n e^{\frac{2\pi}{T} int}}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi i}{T} n\lambda + k}. \tag{2.16.11}$$

Setting $\widehat{f}_{\overline{n}} = \varrho_{\overline{n}} e^{i\theta_{\overline{n}}}, \varrho_{\overline{n}} \geq 0, \theta_{\overline{n}} \in \mathcal{R}$, the first term becomes

$$2\varrho_{\overline{n}} \frac{\sin(\frac{2\pi}{T_0} t + \delta_{\overline{n}})}{2\pi\lambda/T_0}, \tag{2.16.12}$$

while the series in Eq. (2.16.11) can be bounded above uniformly in λ by

$$\sum_{|n|\neq\overline{n}} \frac{|\widetilde{f}_n|}{|-m(\frac{2\pi}{T})^2 n^2 + k|}. \tag{2.16.13}$$

So if $T = \overline{n} T_0$, for some integer $\overline{n}$, and if the force f is arbitrarily small but such that $\widehat{f}_{\overline{n}} \neq 0$, the motion impressed by f to the oscillator may attain an enormous amplitude, as Eqs. (2.16.11)-(2.16.13) show, for small λ.

If T/T_0 is not integer but almost such, $(T/T_0 \simeq \overline{n} \in \mathcal{Z})$, it will happen that the series of Eq. (2.16.10) will contain terms (those with $n = \pm\overline{n}$) with denominators which, even though not vanishing as $\lambda \to 0$, will become very small producing two contributions to Eq. (2.16.10) that could "dominate" the others.

(3) It should be stressed that resonance manifests itself only when the terms $\lambda e^{\alpha_\pm t}$ in Eq. (2.16.8) are small and, therefore, only if $\lambda t/2m \gg 1$. Hence,

although it is true that in the resonating linear oscillator ($T = \overline{n}T_0, \overline{n} \in \mathcal{Z}_+$) a very small force can produce huge oscillations (proportional to λ^{-1}), it is also true that the time it takes for this to happen is very large (proportional to λ^{-1}). Note also that

$$A_+ = \overline{A}_- = \frac{(v_0 - \dot{\overline{x}}_\lambda(0)) - \alpha_-(x_0 - \overline{x}_\lambda(0))}{\alpha_+ - \alpha_-} \tag{2.16.14}$$

becomes singular as $\lambda \to 0$, in resonance cases, because such are $\overline{x}_\lambda(0), \dot{\overline{x}}_\lambda(0)$ [see Eqs. (2.16.10) and (2.16.12)].

(4) Equations (2.16.6) and (2.16.7) give the most general solutions to Eq. (2.16.3) as A_0, φ_0 vary arbitrarily. They show that when $\lambda = 0$, the linear oscillator motions are not longer periodic but, rather, are "sums" of two periodic motions with respective periods T_0 and T equal to the "proper" period of the free oscillator and to the period of the forcing force provided that T/T_0 is not an integer. If T/T_0 is integer, and if the harmonic component of order T/T_0 of the force f does not vanish, the asymptotic motion is even unbounded: "undamped resonance".

Furthermore, in every case, the asymptotic motion depends on the initial datum (through A_0, φ_0). It is clear that the initial datum dependence surviving in the asymptotic regime is due to the absence of friction: analytically this appears via the fact that $e^{\alpha_\pm t} \not\to 0$, since $\mathcal{R}\text{e}\ \alpha_\pm = 0$ if $\lambda = 0$.

(5) The proof of Proposition 23 is a simple discussion of the limit as $\lambda \to 0$ of the expressions (2.16.8) and (2.16.14). No problem arises in the absence of resonance. In the resonant case, the limit is most conveniently discussed by collecting together the first two terms in Eq. (2.16.8) and the two resonant terms in the series (2.16.8) (i.e., those with $n = \pm T/T_0$). The calculations are straightforward and are left to the reader.[9].

2.16.1 Exercises and Problems

1. Determine up to 20%, the asymptotic amplitude of the oscillations of the motions of $\ddot{x} + \lambda\dot{x} + x = f(t)$, $f(t) = (1 - \cos 2\pi t/T)^{-1}$ for $T = 1, 4\pi, \sqrt{2}$. Which, in each case, are the resonant harmonics? (Call "resonance" a harmonic of order $n \in \mathcal{Z}$ if the function $\xi \to (\frac{2\pi}{T})^2\xi^2 - (\frac{2\pi}{T_0})^2$ takes its minimum between n and $n+1$.)

2. Determine the asymptotic amplitude of the motion described by $\ddot{x} + x = f(t)$ with f given as in Problem 1.

3. Estimate how small λ has to be taken so that the amplitude of the asymptotic oscillations described by $\ddot{x} + \lambda\dot{x} + x = f(t)$, with $f(t) = 10^{-3}(1 - 10^{-2}\cos t)^{-1}$, is not smaller than $A = 1, 10, 10^2, 10^6$.

4. Same as Problem 3 with $f(t) = 10^{-3}(1 - 0.99\cos t)^{-1}$.

---

[9] Note that (i) would also directly follow from the regularity theorem (Proposition 3,p. 22, and problem 17, p. 32)

5. Write a computer program for the empirical solution (i.e., without error estimates) of the equation in Problem 1 with the purpose of drawing graphs, in the data space, of the trajectories corresponding to the various choices of the initial datum (using the computer screen and always avoiding tabulation of results).

2.17 An Application: Construction of a Rigorously Periodic Oscillator in the Presence of Friction. The Anchor Escapement, Feedback Phenomena

Che l'una parte l'altra tira e urge
Tin tin sonando con sì dolce nota
Che 'l ben disposto spirto d'amor turge [10]

In §2.12 we saw that a damped harmonic oscillator can move exactly periodically with a period equal to that of the forcing term. Furthermore, any of its motions differs from this periodic one by an amount which becomes exponentially small as $t \to +\infty$. It is natural to try to use this property to build a clock, i.e., a mechanism moving in a rigorously periodic fashion despite friction. However, the difficulty of producing a rigorously periodic force seems to be, at least, of the same order of magnitude as that of producing a periodic motion.

The anchor escapement is a contrivance in a timepiece which controls the motion of the train of wheel work and through which the energy of the weight is delivered to the pendulum by means of impulses which keep the latter in vibration (see: Webster).

This mechanism simultaneously solves the two problems of building a rigorously periodic force and of inducing a rigorously periodic motion. It takes advantage of the presence of friction to cause the oscillator to move asymptotically in a periodic way in the sense that the difference between the actual oscillator's position $x(t)$, at time t, and the position of a certain ideal periodic motion $x_{per}(t)$ tends exponentially to zero as $t \to +\infty$. A very schematic empirical description of the anchor escapement is the following.

The "anchor" is a device set in motion by the oscillator as it passes through the point $x_0 = 0$, for instance, with positive velocity. At this instant, a notched wheel connected to a weight is liberated from a brake and starts moving. A little later, the notch of the wheel reaches the oscillator and accompanies it for a short while, exerting a push on it. Then the notched wheel loses contact with the oscillator, which remains free, allowing the wheel to return

[10] In basic English:
That every part pulls another
tin tin singing, so sweetly:
that the well inclined spirit is filled with love.
(Dante, Paradiso, Canto X).

to its original position by continuing its rotation. In this simplified scheme, the wheel has just one notch instead of the usual few dozens. In the meantime, the oscillator, now free, continues its (damped) oscillation, and the entire process starts afresh at the new passage through x_0 with positive speed.

An attempt to schematize the just-described mechanical system is a motion governed by the equation:

$$m\ddot{x} + \lambda\dot{x} + kx = 0, \qquad x < 0, \tag{2.17.1}$$

$$m\ddot{x} + \lambda\dot{x} + kx = f(\dot{x}_0, \tau) \qquad x \geq 0 \tag{2.17.2}$$

where the action of the notched wheel is schematized by a force $f(\dot{x}_0, \tau)$ depending upon the velocity $\dot{x}_0$ of the last passage through $x_0 = 0$ with positive speed and upon the time τ elapsed since then.

Note that Eqs. (2.17.1) and (2.17.2) are equations of motion quite different from the ones considered in the preceding sections. The force appearing in Eq. (2.17.2) not only depends on time but also upon the past history of the motion itself. Therefore it is not a differential equation in the sense of §2.2, Definition 1. Consequently, we do not even know yet whether Eqs. (2.17.1) and (2.17.2) have a solution, i.e., a C^∞ function $t \to x(t)$, $t \in \mathcal{R}_+$, which turns Eqs. (2.17.1) and (2.17.2) into an identity (not even if f is a C^∞ function of its arguments).

To study Eqs. (2.17.1) and (2.17.2), it is useful to place some restrictions on the form of f which we intend to consider: i.e., it is useful to further specialize the model. This is done to avoid problems too complex from a technical point of view, as well as to avoid developing a theory for too general an f, which may not correspond to a force law that is reasonable for our problem.

For the sake of example, let us assume that $f(\dot{x}_0, \tau)$ vanishes whenever $\dot{x}_0$ does not belong to an interval $[v_-, v_+]$, $0 < v_- < v_+$:

$$\dot{x}_0 \notin [v_-, v_+] \;\to\; f(\dot{x}_0, \tau) = 0. \tag{2.17.3}$$

The assumption corresponds to the fact that when the oscillator sweeps through $x_0 = 0$ too fast, it is never reached by the wheel's notch; while if it sweeps too slowly, the amplitude of oscillation is too small to allow the oscillator to touch the notch.

Assume, also, that once $\dot{x}_0 \in [v_-, v_+]$, the force on the oscillator only depends on the time τ elapsed since the last passage through $x_0 = 0$ with positive speed; i.e.,

$$f(\dot{x}_0, \tau) = P\,\chi(\dot{x}_0)\,g(\tau) \tag{2.17.4}$$

where $\chi(\dot{x}_0) = 1$ if $\dot{x}_0 \in [v_-, v_+]$, and $\chi(\dot{x}_0) = 0$ otherwise, and $t \to g(t) \geq 0$ is a $C^\infty(\mathcal{R})$ function vanishing outside an interval $[a, T_g]$, $a > 0$, $T_g > 0$ with a maximum equal to 1. The constant P, which we shall take as a positive adjustable parameter, models the "intensity" of the force. Physically, one can

imagine that P depends on the weight moving the notched wheel, while g is a detailed description of the wheel action.

Therefore, Eq. (2.17.4) will be considered a mathematical model of the force generated by the anchor escapement. Such a model is only a schematization, where some of the properties of any real mechanism are certainly oversimplified. Nevertheless, as it will be shown, it is a model presenting some interesting characteristics such as, primarily, the "self-control" or "feedback" mechanism providing that the system (2.17.1), (2.17.2), and (2.17.4) "searches automatically", in certain circumstances, for a situation of motion that allows it to move periodically. The function g has a graph like that in Fig.2.6.

Some further properties which we must impose on g should be that g vanishes for $\tau < \alpha$, for some $\alpha > 0$, or for $T > T_g > \alpha > 0$, and T_g should be small compared to the time necessary for an elongation of the oscillator from the position $x_0 = 0$ to the position of maximum distance from $x_0 = 0$.

The time $\alpha > 0$ is a mechanical constant representing the delay between the beginning of the wheel motion and the actual oscillator-notch contact. The $\dot{x}_0$ independence of α is a strong idealization.

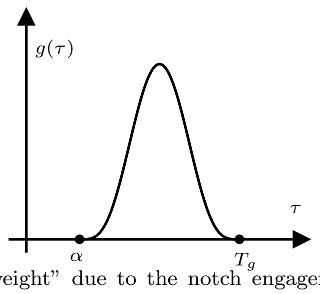


Fig.2.6: The force "per unit weight" due to the notch engagement of the oscillator as a function of time elapsed since the last sweep through the origin.

The physically obvious requirement on T_g can be translated into mathematical terms by requiring $T_g \ll T_0/4$, where $T_0 = 2\pi\sqrt{m/k}$ is the ideal oscillator period. This attempts to translate the fact that the notch has to detach itself from the oscillator before the latter starts swinging back toward the origin. Empirically, the condition $T_g \ll T_0/4$ should guarantee this fact, at least if the friction is small so that it produces negligible effects for times of the order of T_0; i.e., as we saw in §2.12, if $\lambda^2 < 4mk$.

As a conclusion to the above considerations, assume as a model for the anchor escapement Eqs. (2.17.1), (2.17.2), and (2.17.4) with g as in Fig.2.6 with $0 < \alpha < T_g < T_0/4$ and with $\lambda^2 2 < 4mk$, and let us prove th following proposition, which begins the theory of the model.

24 Proposition. *Under suitable compatibility conditions between the parameters P, v_-, v_+ , Eqs. (2.17.1), (2.17.2), and (2.17.4) admit a periodic C^∞ solution defined for $t \in \mathcal{R}_+$.*

Observation. In the upcoming section we shall discuss the compatibility conditions by showing that they can be satisfied at least when λ is small enough.

Later, in §2.19 we shall also show that when λ is small enough and the compatibility conditions are fulfilled, the motions with initial data close enough to those of the periodic motion become close to such motion exponentially fast (see Figs.2.7 and 2.8).

Proof. Let $t \to x(t)$ be a given periodic motion with period $T > T_g$ and such that $x(0) = 0$. Then the function defined, for $t \geq 0$, by

$$\varphi(t) = f(v_0, \tau) = P\,\chi(v_0) g(\tau) \tag{2.17.5}$$

where v_0 is the velocity of the given motion at its last passage through the origin, with positive speed and before time t, and τ is the time elapsed since such time, is a C^∞-periodic function of t, provided the time necessary to return to 0 with positive speed is equal to T itself.[11]

Assuming, then, that $t \to x(t)$, $t \in \mathcal{R}_+$, is a C^∞-periodic motion verifying Eqs. (2.17.1), (2.17.2), and (2.17.4) and period T equal to its first return time to the origin with positive speed and assuming that $v_0 = \dot{x}(0) \in [v_-, v_+]$ we shall have, $\forall\, t \geq 0$,

$$m\,\ddot{x}(t) + \lambda \dot{x}(t) + k\,x(t) \;=\; \varphi(t). \tag{2.17.6}$$

Since $\varphi(t) \equiv P\,g(t)$, $\forall\, t \in [0, T_g]$, and if

$$\widehat{g}_n \stackrel{def}{=} \int_0^T g(\tau)\, e^{-\frac{2\pi}{T} in\tau} \frac{d\tau}{T} \equiv \int_0^{T_g} g(\tau)\, e^{-\frac{2\pi}{T} in\tau} \frac{d\tau}{T}, \tag{2.17.7}$$

recalling that $T_g < T$, it follows (see §2.12)

$$x(t) = \;P \sum_{n=-\infty}^{+\infty} \frac{\widehat{g}_n\, e^{\frac{2\pi}{T} int}}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi}{T} in\lambda + k}, \tag{2.17.8}$$

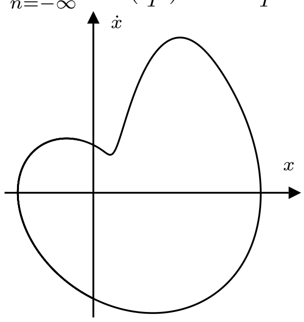


Figure 2.7

Figure 2.7. Graph of a periodic solution $t \to (x(t), \dot{x}(t))$ of Eqs. (2.14.1), (2.14.2), and (2.14.4) with convenient choices of the arbitrary parameters and of the function g.

and the series is uniformly convergent because $\widehat{g}_n$ approaches zero as $n \to \infty$ faster than any power (being the Fourier transform of the C^∞-periodic

[11] It could a priori happen that the motion sweeps through the origin more than twice (even infinitely many times) in an interval of time equal to the period T.

function φ). Of course, we still have to determine T and to check that for such T, Eq. (2.17.8) is really a solution to Eqs. (2.17.1), (2.17.2), and (2.17.4). In other words, we must impose the condition that Eq. (2.17.8) is such that

$$(i) \qquad x(0) = 0; \tag{2.17.9}$$

$$(ii) \qquad \dot{x}(0) \in [v_-, v_+]; \tag{2.17.10}$$

$$(iii) \qquad T > T_g.; \tag{2.17.11}$$

$$(iv)\ T \text{ is the first return time in 0 with positive velocity.} \tag{2.17.12}$$

Relation (2.17.9) is an equation for the period T:

$$0 = \sum_{n=-\infty}^{+\infty} \frac{\widehat{g}_n}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi}{T} in\lambda + k}, \tag{2.17.13}$$

and it should be noted that in this equation, T also appears in the coefficients $\widehat{g}_n$ [see Eq. (2.17.7)].

Then if the parameters v_-, v_+, P, T are such that Eq. (2.17.13) admits at least one solution T and if with this choice of T Eq. (2.14.8) verifies the

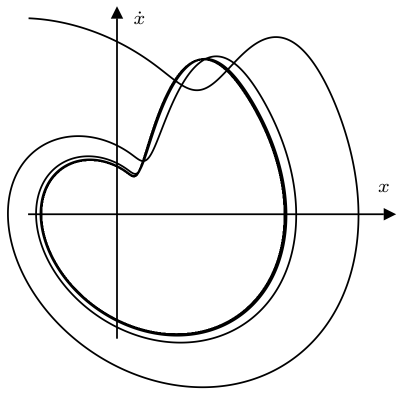


Figure 2.8

Figure 2.8. Graph of a solution $t \to (x(t), \dot{x}(t))$ with initial datum chosen arbitrarily: it becomes indistinguishable from the periodic solution of Fig.2.7 within a few oscillations (three in the precision of the drawing).

compatibility conditions Eq. (2.17.10)-(2.17.12), it follows that Eq. (2.17.8) is a T-periodic solution to Eqs. (2.17.1), (2.17.2), and (2.17.4). mbe

2.17.1 Exercises

1. Choose arbitrarily a function g and $m, k, \lambda, v_-, v_+ > 0$ and write a computer program providing a heuristic (i.e., without error estimate) solution to Eqs. (2.17.1), (2.17.2), and (2.17.4) in which P and the datum $\dot{x}(0)$ are left as free parameters. The output of the program should be a graph like those in Figs. 2.7 and 2.8.

2. Run the above program on a desk computer plotting on the screen the results and finding, by trial and error, a value of P yielding a nontrivial periodic motion.

2.18 Compatibility Conditions for the Anchor Escapement

This section, as well as the next, will suppose some maturity on the reader's part and, therefore, on first reading it would be appropriate to skip the proof of this section and to read the next section only up to the beginning of the proof of Proposition 26.

As promised in the previous section, it will now be shown that if λ is small enough, given v_-, v_+, g with $T_g < T_0/4$, it is possible to fix P so that Eq. (2.17.8) actually verifies the four compatibility conditions Eq. (2.17.9)-2.17.12) and, therefore, is a periodic solution to Eqs. (2.17.1), (2.17.2), and (2.17.4), i.e., to the equation for the anchor-escapement model.

Consider Eq. (2.17.13) as an equation for T^{-1} parameterized by $\lambda > 0$, and let us find some of its solutions having the form

$$T^{-1} = T_0^{-1}\,(1+\lambda\beta) \tag{2.18.1}$$

suggested by the idea that for small λ the oscillator may oscillate with a periodic motion with period close to the period $T_0 = 2\pi\sqrt{m/k}$ of the frictionless oscillator. The equation for T, Eq. (2.17.13), then becomes, after explicitly separating out of the series sum the two complex conjugate terms with $n = \pm 1$:

$$\begin{aligned} 0 =&2\,\mathcal{R}\mathrm{e}\,\Big\{\frac{\widehat{g}_1}{-m(\frac{2\pi}{T_0})^2(1+\lambda\beta)^2+\frac{2\pi}{T_0}i\lambda\,(1+\lambda\beta)+k}\Big\}\\ &+\sum_{\substack{n=-\infty\\ n\neq\pm1}}^{+\infty}\frac{\widehat{g}_n}{-m(\frac{2\pi}{T})^2n^2+\frac{2\pi}{T}in\lambda+k} \end{aligned} \tag{2.18.2}$$

which, using $T_0 = 2\pi\sqrt{m/k}$, becomes

$$\begin{aligned} 0 =&2\mathcal{R}\mathrm{e}\,\Big\{\frac{\widehat{g}_1}{-k(2\beta+\beta^2\lambda)\lambda+\frac{2\pi}{T_0}i\lambda\,(1+\lambda\beta)}\Big\}\\ &+\sum_{\substack{n=-\infty\\ n\neq\pm1}}^{+\infty}\frac{\widehat{g}_n}{-m(\frac{2\pi}{T})^2n^2+\frac{2\pi}{T}in\lambda+k}. \end{aligned} \tag{2.18.3}$$

We see that for small λ the first of the above two addends shows a small denominator. To avoid having to study an equation with small denominators, multiply Eq. (2.18.3) by λ. Then, $\forall\,(\lambda,\beta)$, $\lambda\beta \geq -\frac{1}{2}$ (so that $T_0^{-1}(1+\lambda\beta) > 0$), it is possible to define

$$\Phi(\lambda,\beta) \stackrel{def}{=} 2\,\mathcal{R}\mathrm{e}\,\left\{\frac{\widehat{g}_1}{-k(2\beta+\beta^2\lambda)\lambda+\frac{2\pi}{T_0}i\lambda\,(1+\lambda\beta)}\right\}$$
$$+\,\lambda\sum_{\substack{n=-\infty\\ n\neq\pm1}}^{+\infty}\frac{\widehat{g}_n}{-m(\frac{2\pi}{T})^2n^2+\frac{2\pi}{T}in\lambda+k}. \tag{2.18.4}$$

where $T^{-1}=T_0^{-1}\,(1+\lambda\beta)$ and it is perhaps worth recalling that $\widehat{g}_n$ are also T dependent. Rewrite Eq. (2.18.3) as

$$\Phi(\lambda,\beta)=0 \tag{2.18.5}$$

with the additional restriction $\lambda\beta>-\frac{1}{2}$ [to be amply sure that the denominators in Eq. (2.18.4) do not vanish]. To study Eq. (2.18.5), note that the equation $\Phi(0,\beta)=0$ leads to

$$2\,\mathcal{R}\mathrm{e}\,\frac{\widehat{g}_1}{-2k\beta+\frac{2\pi}{T_0}i}=0 \tag{2.18.6}$$

which, defining

$$b(T^{-1})\stackrel{def}{=}\mathcal{R}\mathrm{e}\,\widehat{g}_1,\qquad c(T^{-1})\stackrel{def}{=}\mathcal{I}\mathrm{m}\,\widehat{g}_1, \tag{2.18.7}$$

has as a solution the quantity β_0:

$$\beta_0=\frac{1}{2k}\frac{2\pi}{T_0}\frac{c(T_0^{-1})}{b(T_0^{-1})}, \tag{2.18.8}$$

provided that $b(T_0^{-1})\neq 0$. Note also that $b(T_0^{-1})\neq 0$ as it follows from [see Eq. (2.17.7)]

$$b(T_0^{-1})=\int_0^{T_g}\frac{d\tau}{T_0}\,g(\tau)\cos\frac{2\pi}{T_0}\tau,\qquad c(T_0^{-1})=\int_0^{T_g}\frac{d\tau}{T_0}\,g(\tau)\sin\frac{2\pi}{T_0}\tau, \tag{2.18.9}$$

thanks to the assumption $T_g<T_0/4$ which implies that for $T\in(0,T_g)$, the sine and cosine in Eq. (2.18.9) are positive.

The above remarks give hope for the existence of a solution to Eq. (2.18.5) having the form

$$\beta(\lambda) = \beta_0 + O(\lambda), \tag{2.18.10}$$

at least if λ is small. If this were true, the velocity $\dot{x}(0)$ could be computed from Eq. (2.17.8):

$$\begin{aligned}\dot{x}(0) =& P \sum_{n=-\infty}^{+\infty} \frac{2\pi in}{T} \frac{\widehat{g}_n}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi}{T} in\lambda + k},\\ =&\frac{P}{\lambda}\Big[2\mathcal{R}\mathrm{e}\ \frac{\widehat{g}_1}{-k(2\beta+\lambda\beta^2)+\frac{2\pi}{T}i}\\ &+\sum_{\substack{n=-\infty\\ n\neq\pm1}}^{+\infty} \frac{2\pi i}{T} n \frac{\widehat{g}_n}{-m(\frac{2\pi}{T})^2 n^2 + \frac{2\pi}{T} in\lambda + k}\Big].\end{aligned} \tag{2.18.11}$$

Hence, Eqs. (2.18.11) and (2.18.10) would imply, with some algebra (and patience),

$$\dot{x}(0) =\frac{P}{\lambda}\Big[2\mathcal{R}\mathrm{e}\ \{\frac{2\pi}{T_0}\frac{ib(T_0^{-1}) - c(T_0^{-1})}{-2k\beta_0 + \frac{2\pi i}{T_0}}\} + O(\lambda)\Big] = \frac{2P}{\lambda}\big(b(T_0^{-1}) + O(\lambda)\big), \tag{2.18.12}$$

having used Eq. (2.18.8).

Therefore, if λ is so small that in Eq. (2.18.12) $|O(\lambda)| < b(T_0^{-1})$, it is $\dot{x}(0) \neq 0$, and P can be so chosen that $\dot{x}(0) \in [v_-, v_+]$. Note that $P \to 0$ proportionally to λ, as $\lambda \to 0$, if one imposes $\dot{x}(0) \in [v_-, v_+]$: this agrees with the obvious empirical observation that the "weight" necessary to move the oscillator must be small in proportion to friction.

Similarly, starting from Eqs. (2.18.7) and (2.18.10), one could check Eq. (2.17.12) for small λ. It can in fact be seen that it is enough to verify Eq. (2.17.12) by replacing $x(t)$ in Eq. (2.17.8) with the only contributions to the series (2.17.8) coming from the $n = \pm 1$ terms (which in the preceding discussion seem to be the only important ones for small λ, as far as the computation of T and of $\dot{x}(0)$ are concerned). For such an approximation to $x(t)$, the statement of Eq. (2.17.12) is, however, obvious since such an approximate motion is a harmonic motion with period T. Elaboration of the details is left to the reader.

Finally, Eq. (2.17.11) would also immediately follow from Eqs. (2.18.1) and (2.18.10) for small λ.

The above analysis can be summarized in the following proposition.

25 Proposition. *If Eq. (2.18.5), as an equation for β parameterized by λ, admits a solution having the form of Eq. (2.18.10) for λ small enough, then*

the equation for the anchor-escapement model [Eqs. (2.17.1), (2.17.2), and (2.17.4)] admits a periodic solution with period T such that:

$$T^{-1} = T_0^{-1}(1 + \beta_0\lambda + O(\lambda)) \tag{2.18.13}$$

if λ is sufficiently small and if P is suitably chosen.

Therefore, to complete the solution of our question, it only remains to verity that Eq. (2.18.5) does indeed admit a solution β like Eq. (2.18.10) for small λ.

A pair (λ, β) verifying Eq. (2.18.5) is already known, namely the pair $(0, \beta_0)$; hence it is natural to try to treat Eq. (2.18.5) through the implicit function theorem (see Appendix G). By this theorem, it will be enough to check that the function Φ, Eq. (2.18.4), defined in the open set of $\mathcal{R}^2$ containing the points (λ, β) such that $\lambda\beta > -\frac{1}{2}$, is of class C^∞ in its domain of definition and has a first-order derivative with respect to β such that

$$\frac{\partial \Phi}{\partial \beta}(0, \beta_0) \neq 0. \tag{2.18.14}$$

In this case, Eq. (2.18.5) will admit a solution β for λ small enough, like

$$\beta = \beta_0 - \frac{(\partial\Phi/\partial\lambda)(0, \beta_0)}{(\partial\Phi/\partial\beta)(0, \beta_0)}\,\lambda + o(\lambda) \tag{2.18.15}$$

To see that Φ is a C^∞ function near $(0, \beta_0)$, one shows that from expression in Eq. (2.17.7) and from estimates in Eq. (2.13.7) it follows that

$$\widehat{g}_n = \frac{1}{\big((2\pi i/T)n\big)^k} \int_0^{T_g} \frac{d^k g}{d\tau^k}(\tau)\, e^{-\frac{2\pi i}{T} n\tau} \frac{d\tau}{T} \tag{2.18.16}$$

and, by Newton's formula for the p-th derivative of a product:

$$\begin{aligned} \frac{\partial^p \widehat{g}_n}{\partial (T^{-1})^p} =& \frac{1}{(2\pi i n)^k} \int_0^{T_g} d\tau\, \frac{d^k g}{d\tau^k}(\tau)\, e^{-\frac{2\pi}{T} i n\tau} \\ & \cdot \sum_{j=0}^{p} (-2pni)^{p-j} (-k+1)(-k)\ldots(-k-j+2)(T^{-1})^{-k+1-j}; \end{aligned} \tag{2.18.17}$$

hence

$$\Big|\frac{\partial^p \widehat{g}_n}{\partial (T^{-1})^p}\Big| \leq \max_{\substack{0 \leq \tau \leq T_g \\ 0 \leq j \leq p}} \Big[\frac{(k+p)^j (2\pi |n| T_g)^{p-j}}{T^{k-1+j}} \Big|\frac{d^k g}{d\tau^k}(\tau)\Big|\Big] \tag{2.18.18}$$

which implies that as long as $T < +\infty$ (i.e., $1 + \beta\lambda > 0$), the function in Eq. (2.18.4) is a C^∞ function of β and λ.

The last three relations also imply that the derivatives of Φ with respect to λ and P can be computed by term-by-term differentiation of the series

defining Φ in the region $\lambda\beta > -\frac{1}{2}$. After a brief computation, such a term-by-term differentiation evaluated at $(0, \beta_0)$ yields:

$$\frac{\partial \Phi}{\partial \beta}(0, \beta_0) = \big(\frac{T_0}{\pi}\big)^2 \frac{b(T_0^{-1})^3}{b(T_0^{-1})^2 + c(T_0^{-1})^2} \tag{2.18.19}$$

and this check of Eq. (2.18.14) concludes the discussion of the compatibility conditions showing that they can indeed be satisfied for small enough λ.

2.19 Encore on Anchor Escapement: Stability of the Periodic motion

In the preceding sections we showed that the anchor-escapement model [Eqs. (2.17.1), (2.17.2), and (2.17.4)] admits a periodic solution for small enough friction if the intensity of weight P is suitably chosen.

Imagining to have fixed P conveniently in terms of λ, such a periodic motion will be denoted $t \to \overline{x}(t)$, $t \in \mathcal{R}_+$. However existence of the motion $\overline{x}$ is not interesting in itself for applications. In fact, to put the system in this state of motion, one would have to impress exactly the velocity v_0 at $t = 0$, with v_0 defined by Eq. (2.18.11) after putting the oscillator in $x_0 = 0$. In fact, these are the initial data of the periodic motion corresponding to the a priori given λ and P.[12]

The periodic motion studied in the preceding sections is interesting for applications only if it is "stable", i.e., only if starting the system in an initial state $x(0) = 0, \dot{x}(0) = v_0 + \eta$, perturbed with respect to that which would generate a periodic motion, would produce a motion $t \to x_\eta(t)$, $t \in \mathcal{R}_+$, according to Eqs. (2.17.1), (2.17.2), and (2.17.4), which exists and is unique, at least for small η, and, furthermore,

$$|x_\eta(t) - \overline{x}(t - \tau_\eta)| \xrightarrow[t\to+\infty]{} 0 \tag{2.19.1}$$

if τ_η is suitably chosen.

In applications, one would like to require more: for instance, one would wish that the limit (2.19.1) is attained with an exponential speed with a halving time of the order of the period T of the periodic motion. In such a case, after a "few" oscillations, the motion would be identical to the rigorously periodic one, for all practical purposes. This is what actually occurs in the pendulum clock.

[12] One should also show that to such an initial datum an actually periodic motion does follow: i.e., one should prove a uniqueness theorem, at least for the initial data under examination. This is possible, as well as it is also possible to show a uniqueness property on the perturbed motions that will be met in this section. However, we shall not enter into the proof of the validity of the uniqueness properties that interest us: the reader should do this as a problem. Note that Proposition 1, p. 14, does not directly apply here, since the equations do not have the form contemplated in §2.2.

To examine the stability problem, the following proposition will be proved.

26 Proposition. *The periodic motion of the anchor-escapement model,* $t \to \overline{x}(t)$, $t \in \mathcal{R}_+$, *built in §2.17 and §2.18, is stable in the sense expressed in Eq. (2.19.1) if* λ *is small enough. The limit (2.19.1) is reached exponentially with a halving time* $T_{1/2}$ *of the order of magnitude*

$$T_{1/2} \simeq \max(T, 2m\lambda^{-1}) \tag{2.19.2}$$

Observations.
(1) During the proof, the role of friction and its importance will clearly appear. It is a rather general rule that the dissipative motions are more stable than the corresponding frictionless motions, as long as the friction is not too strong. The price paid for this stability, of obvious and essential importance in applications, is naturally the necessity of the action of a force to maintain the motion itself.
(2) One could require, and prove, stability with respect to initial data that are more general and realistic than those considered in Proposition 26. For instance, with respect to initial data like $x(0) = \varepsilon, x(0) = v_0 + \eta$, the theory and results would be essentially the same.
(3) Proposition 26 concludes our theory of the anchor escapement. One should clearly bear in mind that the mathematical equations (2.17.1), (2.17.2), and (2.17.4) are just a model, in some respects not very satisfactory. For instance, $\dot{x}_0$ independence of the force $f(\dot{x}_0, \tau)$, once $\dot{x}_0 \in [v_-, v_+]$, is unrealistic.
(4) However, the model considered performs perfectly one of the typical functions of models and clarifies the possibility of the existence of an important mechanism which would also have to be present in more refined models: the possibility of a motion controlling itself via a feedback reaction inducing it to move periodically after a short while. This self-control, understood and practically realized at a time when the field of mechanics was new, is a phenomenon which appears in many models concerning the most diverse physical systems. The design and construction of the most precise machines are based on it, as well as the very possibility of their existence.

Proof. Define [see Eq. (2.19.1) and the preceding lines for notation]:

$$x_\eta(t) = \overline{x}(t) + \xi(t) \tag{2.19.3}$$

and let us show the existence of a C^∞ solution of Eqs. (2.17.1), (2.17.2), and (2.17.4) verifying the initial conditions $x_\eta(0) = 0, \dot{x}_\eta(0) = \dot{\overline{x}}(0) + +\eta$, provided η is small enough and the values of P, λ are such that the periodic motion $t \to \overline{x}(t)$, $t > 0$, exists. Call T the period of $\overline{x}$. First, note that if $t \to \xi_1(t)$ is the solution of the equation

$$m\ddot{\xi}_1 + \lambda\dot{\xi}_1 + k\xi_1 = 0, \quad \xi_1(0) = 0,\ \dot{\xi}_1(0) = \eta, \qquad t \in \mathcal{R}_+ \tag{2.19.4}$$

and if $\overline{T}_1$, is the first positive time when the motion $t \to \overline{x}(t) + \xi_1(t)$ passes through the origin with positive speed, then the motion solves Eqs. (2.17.1), (2.17.2), and (2.17.4) for $\tau \in [0, \overline{T}_1]$ if η is small. To understand this property, note that the solution of Eq. (2.19.4) is

$$\xi_1(t) = \eta\, e^{-\frac{\lambda}{2m}t} \frac{\sin \sqrt{k/m - \lambda^2/4m^2}\, t}{\sqrt{k/m - \lambda^2/4m^2}} \tag{2.19.5}$$

hence, $\forall\, t \geq 0$:

$$\begin{aligned} |\xi_1(t)| \leq & \frac{|\eta|}{\sqrt{k/m - \lambda^2/4m^2}}, \\ |\dot{\xi}_1(t)| \leq & \frac{|\eta|}{\sqrt{k/m - \lambda^2/4m^2}} (1 + \frac{1}{\sqrt{k/m - \lambda^2/4m^2}}) \end{aligned} \tag{2.19.6}$$

Then, if $|\eta|$ is small enough, to fix the ideas $|\eta| < \delta_\lambda$ with δ_λ suitably chosen, it is clear that Eq. (2.19.3) with $\xi_1(t)$ replacing $\xi(t)$ verifies Eqs. (2.17.1), (2.17.2), and (2.17.4).

To estimate a choice for δ_λ, the following conditions must be imposed:
(1) $T_g < \overline{T}_1 < T + \alpha$, $\forall\, |\eta| < \delta_\lambda$;
(2) the velocity at the first passage through the origin is negative and at the second passage is positive.

Such conditions are true for the reference motions $\overline{x}$ if λ is small enough since, in such a case, as already mentioned and used in §2.18, the reference motion is almost a harmonic motion of period $\sim T_0$ for which the conditions under analysis manifestly hold. Therefore, by continuity, they must remain true for the motion $t \to \overline{x}(t) + \xi_1(t)$ if η is small. We leave the elaboration of the details to the reader.

The fact that $t \to \overline{x}(t) + \xi_1(t)$ is a solution for $t \in [0, \overline{T}_1]$ will not, in general, remain true for $t > \overline{T}_1$, because in Eq. (2.17.4) the time τ is now counted beginning at $\overline{T}_1$, and $T \neq \overline{T}_1$, in general.

To study the motions for times following $\overline{T}_1$, define

$$\eta_1 \stackrel{def}{=} \dot{\overline{x}}(\overline{T}_1) + \dot{\xi}_1(\overline{T}_1) - \dot{\overline{x}}(0); \tag{2.19.7}$$

then if $|\eta_1| < \delta_\lambda$, we can define, as we already saw, the function $t \to (x(t) - \overline{x}(t) - \overline{T}_1) = \xi_2(t)$ where $\xi_2(t)$ is defined for t between $\overline{T}_1$, and the first instant $\overline{T}_2$ successive to $\overline{T}_1$, when the motion sweeps through 0 with positive speed for the first time, as the solution of Eq. (2.19.4) with initial datum:

$$\overline{\xi}_2(\overline{T}_1) = 0, \quad \dot{\overline{\xi}}_2(\overline{T}_1) = \eta_1. \tag{2.19.8}$$

Repeating indefinitely the argument, it is possible to define $\eta_2, \eta_3, \ldots$ *provided* $|\eta_i| < \delta_\lambda$, $i = 1, 2, \ldots$, thus obtaining the definition of the times $\overline{T}_1, \overline{T}_2, \overline{T}_3, \ldots$ corresponding to the successive passages through 0 with positive speed.

The stability property asserted in the proposition will have been proven once will have been shown the existence of two constants $c_\lambda > 0$, $0 < \theta_\lambda < 1$ such that for all η, $|\eta| < \delta_\lambda$:

$$|\overline{T}_{p+1} - (\overline{T}_p + T)| \le c_\lambda \, \theta_\lambda^p, \qquad p = 1, 2. \dots \tag{2.19.9}$$

and

$$|\eta_p| \le \theta_\lambda^p |\eta| \tag{2.19.10}$$

at least for small λ. Setting $\overline{T}_0 = 0$, the constant τ_η [see Eq. (2.19.1)] will then be

$$\tau_\eta = \sum_{i=1}^{\infty} (\overline{T}_i - (T + \overline{T}_{i-1})). \tag{2.19.11}$$

Let us then show the validity of Eqs. (2.19.9) and (2.19.10). If $T_0 = 2\pi\sqrt{m/k}$, the value for θ_λ that we shall find will have the form

$$\theta_\lambda = e^{-\frac{\lambda T_0}{2m}} (1 + o(\lambda)). \qquad T_0 = 2\pi \sqrt{\frac{m}{k}} \tag{2.19.12}$$

for λ small enough (note that $\theta_\lambda < 1$ as soon as λ is so small that $(-\lambda T_0/2m + \frac{1}{2}\lambda^2 T_0^2/4m^2 + e^{-\lambda T_0/2m}\, o(\lambda) < 0$, as it is seen by expanding the exponential to second order). This will also prove Eq. (2.19.2) and, neglecting the infinitesimal $o(\lambda)$ in Eq. (2.19.12), it is seen that the larger the friction (compatibly with the supposed $\lambda^2 < 4mk$ and with the existence of $\overline{x}$), the faster the motion tends to become periodic.

To discuss Eqs. (2.19.9) and (2.19.10), one has to find a more concrete expression for $\overline{T}_1$, and, in general, for $\overline{T}_i$, $i \ge 1$. Let $\overline{T}_1 = T + \kappa_1$: the equation for κ_1 is

$$x_\eta(T + \kappa_1) = 0, \tag{2.19.13}$$

with the added condition that $T + \kappa_1$ should be the first positive time when the oscillator passes again through the origin with positive speed.

For $\kappa, \eta \in \mathcal{R}^2$, $\kappa > -T$, define

$$\psi(\kappa, \eta) = \overline{x}(T + \kappa) + \xi_1(T + \kappa) \tag{2.19.14}$$

and Eq. (2.19.13) becomes

$$\psi(\kappa, \eta) = 0. \tag{2.19.15}$$

Since, as seen in the preceding sections, $t \to x(t), t \in \mathcal{R}_+$, is a C^∞ function and, obviously, so is $(\eta, t) \to \overline{x}(t)$, we can say that ψ is a C^∞ function on its domain of definition, $\kappa > -T$.

Furthermore, by Eqs. (2.19.14) and (2.19.5), it is

$$\begin{aligned}
&\psi(0,0)=0, \qquad \text{and} \qquad \frac{\partial\psi}{\partial\kappa}(0,0)=v_0,\\
&\frac{\partial\psi}{\partial\eta}(0,0)=e^{-\frac{\lambda}{2m}T}\,\frac{\sin(k/m-\lambda^2/4m^2)^{\frac12}T}{(k/m-\lambda^2/4m^2)^{\frac12}}
\end{aligned}\tag{2.19.16}$$

Then, by the implicit function theorem (see Appendix G), there is, for small η, a unique small solution of Eq. (2.19.15) which we denote $\kappa_1(\eta)$, and

$$\kappa_1(\eta)=-\frac{e^{-\frac{\lambda}{2m}T}}{v_0}\,\frac{\sin(k/m-\lambda^2/4m^2)^{\frac12}T}{(k/m-\lambda^2/4m^2)^{\frac12}}\,\eta\,+\,o_\lambda(\eta),\tag{2.19.17}$$

where the index λ in $o_\lambda(\eta)$ recalls that the infinitesimal depends also on λ. By taking Eq. (2.18.13) into account:

$$T=T_0\,(1-\beta_0\lambda+o(\lambda)),\tag{2.19.18}$$

and using $\sin\sqrt{\frac{k}{m}}T_0=0$, one finds, with simple steps, from Eqs. (2.19.17) and (2.19.18), that

$$\kappa_1(\eta)=\frac{\beta_0T_0}{v_0}(\lambda+o'(\lambda)+o_\lambda(\eta))\tag{2.19.19}$$

where $o'(\lambda)$ is a suitable infinitesimal of order λ^2.

It then becomes possible to compute η_1:

$$h_1=\dot{\overline{x}}(T+\kappa_1)+\dot\xi_1(T+\kappa_1)-v_0=\ddot{\overline{x}}(T)\kappa_1+\dot\xi_1(T+\kappa_1)+\widetilde{o}_\lambda(\kappa_1)\tag{2.19.20}$$

where $\widetilde{o}_\lambda(\kappa_1)$, is a λ-dependent second-order infinitesimal: this expression arises just by expanding $\dot{\overline{x}}$ in Taylor series near T and using $\dot{\overline{x}}(T)=v_0$.

The equations of motion (2.17.1) imply that $\ddot{\overline{x}}(T)=\ddot{\overline{x}}(0)=-\frac{\lambda}{m}v_0$ and Eq. (2.19.20) implies [via Eqs. (2.19.18) and (2.19.19) and some patience]:

$$\eta_1=\eta\,e^{-\frac{\lambda T_0}{2m}}(1+\widetilde{o}(\lambda)+\widetilde{O}_\lambda(\eta)),\tag{2.19.21}$$

where $\widetilde{o}$ is an infinitesimal of higher order in λ while $\widetilde{O}_\lambda(\eta)$ is a λ-dependent infinitesimal of the same order as η. Therefore there is a $\delta'_\lambda<\delta_\lambda$ sufficiently small so that $|\widetilde{O}_\lambda(\eta)|\le|\widetilde{o}(\lambda)|$, $\forall\,|\eta|<\delta'_\lambda$; then Eq. (2.19.21) implies that

$$|\eta_1|\le\theta_\lambda\,|\eta|,\qquad\forall\,|\eta|<\delta'_\lambda\tag{2.19.22}$$

with θ_λ, given by Eq. (2.19.12).

Hence, if λ is small enough one finds that $|\eta|<\delta'_\lambda$ implies $|\eta_1|<\delta'_\lambda$ and the argument can be indefinitely repeated to estimate successively $|\eta_1|,|\eta_2|,\ldots$. Then from Eqs. (2.19.22) and (2.19.19), the Eqs. (2.19.9) and (2.19.10) follow, and Proposition 26 is thus proved. mbe

2.19.1 Problems

1. Investigate heuristically the stability of the solutions of Eqs. (2.17.1), (2.17.2), and (2.17.4), using the computer program of problem 1, §2.17. For each value of λ let v_0, x_0 be the data, at time zero, of a periodic motion verifying Eqs. (2.17.1), (2.17.2) and (2.17.4); let the computer draw on the screen the graph of the periodic motion superimposed with the graph of the motion of a harmonic oscillator with the same mass and elastic constant (but no friction nor forcing term). Repeat this operation as λ varies using it to compare visually the two motions.

2. Same as Problem 1, replacing kx by $k \sin x$ in Eqs. (2.17.1) and (2.17.2) (i.e., replacing the basic oscillator with a pendulum).

2.20 Frictionless Forced Oscillations: Quasi-Periodic Motions

In §2.16 it has been shown that, under the action of a periodic force, a frictionless harmonic oscillator moves with a motion "sum" (or "superposition") of two periodic motions with respective periods equal to the proper oscillator period T_0 and to the forcing term period T, provided T/T_0 is not an integer.

The proposition in this section will help to visualize some remarkable properties of such motions. One of them appears by representing them as motions on the data space (see §2.6), i.e., on the plane $\mathcal{R}^2$ thought of as the space of the initial velocities and positions. This means that the motion $t \to x(t)$, $t \in \mathcal{R}_+$, solution of Eq. (2.16.3), i.e. of $m\,\ddot{x} + k\,x = f$, is represented by a curve $t \to (\dot{x}(t), x(t)), t \in \mathcal{R}_+$. This is a representation of the motion which we have not yet used: it is somewhat redundant because once $t \to x(t)$ is given, its t-derivative is automatically given. On the other hand, every point of the curve $t \to (\dot{x}(t), x(t)), t \in \mathcal{R}_+$, completely determines the motion. Also it may sometimes be useful to know which are the pairs $(\dot{x}, x)$ which can appear during the evolution of a given motion. In such a case, this information can be directly extracted from the geometric locus described in $\mathcal{R}^2$ by $t \to (\dot{x}(t), x(t)), t \in \mathcal{R}_+$, without having to know explicitly which values of t correspond to the various points of the locus.

Therefore, in the data space, a periodic motions appears as a closed curve. A motion like those met in §2.12, asymptotically periodic, appears as a curve spiraling around the closed curve representing the periodic motion and becoming indefinitely closer to it.

The structure of a superposition of two periodic motions in the data space representation is of particular interest: it is elucidated by the following well-known proposition (Euler theorem).

27 Proposition. *Let $f, g \in C^\infty(\mathcal{R})$ be two periodic functions with minimal period 2π and let f', g' be their first derivatives. Given $\omega, \omega_0 > 0$, consider the motion in $\mathcal{R}^2$ described by $t \to (\eta(t), \xi(t))$:*

$$\eta(t) = \omega f'(\omega t) + \omega_0 g'(\omega_0 t), \qquad \xi(t) = f(\omega t) + g(\omega_0 t) \tag{2.20.1}$$

Such a motion[13] *is periodic if and only if* ω/ω_0 *is a rational number. If* ω/ω_0 *is irrational, the curve* $t \to (\eta(t), \xi(t))$, $t > t_0$, $\forall\, t_0 \in \mathcal{R}_+$ *densely fills the region* $\Omega_{gf,g}$*:*

$$\begin{aligned}\Omega_{f,g} =& \{(\eta,\xi) \,|\, (\eta,\xi) \in \mathcal{R}^2 : \\ & \eta = \omega f'(\alpha) + \omega_0 g'(\beta), \xi = f(\alpha) + g(\beta),\ \alpha, \beta \in [0, 2\pi]\}.\end{aligned} \tag{2.20.2}$$

Observations.

(1) The region $\Omega_{f,g}$ can be easily visualized. Consider the curve Γ_f in the (η, ξ) plane, having equations

$$\eta = \omega f'(\alpha), \quad \xi = f(\alpha), \qquad \alpha \in [0, 2\pi] \tag{2.20.3}$$

By the periodicity of f, this is a closed curve [see Fig. 2.9]. Given $\alpha \in [0, 2\pi]$, consider the curve $\Gamma_g(\alpha)$ with equations

$$\eta = \omega f'(\alpha) + \omega_0 g'(\beta), \quad \xi = f(\alpha) + g(\beta), \qquad \beta \in [0, 2\pi] \tag{2.20.4}$$

which, since g, too, is periodic, is a closed curve "around" $(\omega f'(\alpha), f(\alpha))$.

As α varies in $[0, 2\pi]$, the curve $\Gamma_g(\alpha)$ "glides along Φ_f and "sweeps" the region $\Omega_{f,g}$. A simple case is illustrated in Fig. 2.9.

(2) The relevance of this proposition for the harmonic non resonant forced oscillations is obvious after the discussion of §2.16 (see (iii) in

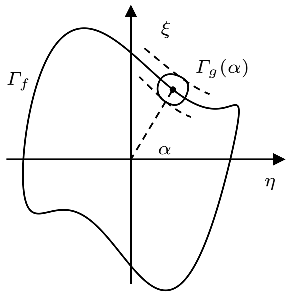


Fig.2.9.: The region swept densely by the quasi periodic motion with irrational ratio of the periods is the region swept by the curve $\Gamma_g(\alpha)$ as α varies.

Proposition 23). It shows that such oscillations, when T and T_0 have an irrational ratio, are not periodic although they come back as close as desired to the initial datum, provided one waits long enough.

(3) Also, for the purpose of future applications, it is interesting to give a geometric interpretation to Proposition 27 when ω/ω_0 is irrational. In this

[13] Remark that $\eta(t) = \dot{\xi}(t)$.

case, the analytic expression of the trajectory density in $\Omega_{f.g}$ is: given $\sigma > 0$ and $t_0 \in \mathcal{R}_+$, for all $(\alpha, \beta) \in [0, 2\pi]$, there is $t_\sigma(\alpha, \beta) > t_0$ to such that

$$|(\alpha - \omega t_\sigma(\alpha, \beta)) \mathrm{mod}\, 2\pi| < \sigma, |(\beta - \omega_0 t_\sigma(\alpha, \beta)) \mathrm{mod}\, 2\pi| < \sigma, \tag{2.20.5}$$

i.e., there are two integers $m_\sigma(\alpha, \beta)$ and $n_\sigma(\alpha, \beta)$ such that

$$|\alpha - \omega t_\sigma(\alpha, \beta) - 2\pi m_\sigma(\alpha, \beta)| < \sigma, \quad |\beta - \omega_0 t_\sigma(\alpha, \beta) - 2\pi n_\sigma(\alpha, \beta)| < \sigma, \tag{2.20.6}$$

Now think of the plane $\mathcal{R}^2$ as being paved with squares with side size 2π and with corners at $(2\pi r, 2\pi s)$, r and s being integers. In this plane, consider the straight line through the origin with slope ω_0/ω:

$$y = \omega_0 t, \; x = \omega t, \qquad t \in \mathcal{R} \tag{2.20.7}$$

and the half-line corresponding to $t \geq t_0$.

Next, identify the points of the plane whose coordinates differ by integer multiples of 2π (see Fig. 2.10). The just-described line can now be thought of as a set of segments in the square $[0, 2\pi]^2$, where corresponding points on opposite sites are identified (topologically, we can say that we regard the square $[0, 2\pi]^2$ as a two dimensional torus).

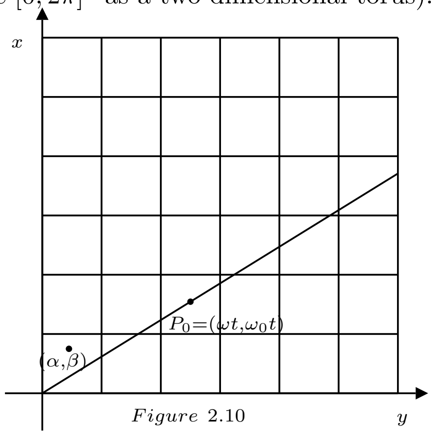


Figure 2.10

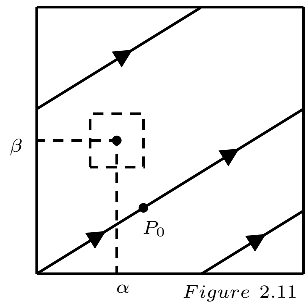


Figure 2.11

The figures represent a quasi periodic motion in the plane and its image on the torus

Equation (2.20.6) says that at least one of the segments associated with the line of Eq. (2.20.7) with $t \geq t_0$ cuts the square neighborhood of side 2σ around (α, β) (see Fig. 2.11).

In other words, the half-line of Eq. (2.20.7) with $t \geq t_0$, brought back inside $[0, 2\pi]^2$ through the identification of the points of the plane mod 2π (i.e., thought of as a coil around the torus) densely fills $[0, 2\pi]^2$.

Proof. If $\omega/\omega_0 \equiv T_0/T = p/q =$ (ratio of two relatively prime integers), where $T \overset{def}{=} 2\pi/\omega, T_0 \overset{def}{=} 2\pi/\omega_0$, then the motion of Eq. (2.20.1) is periodic

with period $T' = pT = qT_0$. As an exercise, the reader can show that in the geometric interpretation of Fig. 2.11, this means that the line becomes a finite set of segments (forming a closed curve if $[0, 2\pi]^2$ is thought of as a torus).

Suppose now that ω/ω_0 is irrational. Define for every integer n the number τ_n as

$$\alpha - \omega\tau_n + 2\pi n = 0 \qquad \longleftrightarrow \qquad \tau_n = \frac{\alpha + 2\pi n}{\omega} \tag{2.20.8}$$

To check Eq. (2.20.5) and, therefore, the validity of the Proposition, it will suffice to show that given $n_0 \in \mathcal{Z}$ and $\sigma > 0$ arbitrarily, there exists $n \in \mathcal{Z}$, $n \geq n_0$, and $m(n) \in \mathcal{Z}$ such that

$$|\beta - \omega_0\tau_n - 2\pi m(n)| < \sigma. \tag{2.20.9}$$

It is useful for the reader to understand (along the lines of observation 3) the geometrical meaning of Eqs. (2.20.8) and (2.20.9) (exercise).

By substituting τ_n, given by Eq. (2.20.8), in Eq. (2.20.9) one finds

$$|\beta - \frac{\omega_0}{\omega}\alpha - 2\pi\frac{\omega_0}{\omega}n - 2\pi m(n)| < \sigma, \tag{2.20.10}$$

i.e. setting $\varphi_0 = \beta - \frac{\omega_0}{\omega}\alpha$:

$$|\varphi_0 - 2\pi\frac{\omega_0}{\omega}n - 2\pi m(n)| < \sigma \tag{2.20.11}$$

Eq. (2.20.11) has a geometric interpretation which is convenient to illustrate: consider the unit circle and its rotation R by an angle $\theta = 2\pi(\omega_0/\omega)$ (see Fig. 2.12). The point with angular coordinate $2\pi(\omega_0/\omega)n$ can be interpreted as the image of a point 0 on the circle under the action of the rotation R^n, i.e., of n successive rotations R. If φ_0 is also interpreted as a point on the circle, Eq. (2.20.11) means that the rotation R^n brings the origin to an angular distance from φ_0 less than σ.

Then our problem is to show the existence, given $\sigma > 0$, of infinitely many integers $n > 0$ such that the rotation R^n brings 0 to an angular

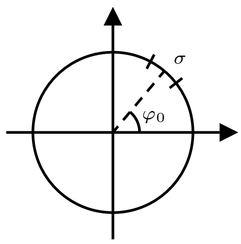


Figure 2.12

distance of less than σ from φ_0. In order to show this, it will be enough to show that there is $\widetilde{n} > 0$ such that $R^{\widetilde{n}}$ displaces the point 0 by a non vanishing quantity ε with modulus less than σ. In fact, when this happens, it is manifest that with a rotation R^n, $n = k\widetilde{n}$, $k = 0, 1, 2, \ldots$, one successively displaces 0 by $\varepsilon, 2\varepsilon, 3\varepsilon, \ldots, k\varepsilon$, and therefore, sooner or later (and infinitely often), one

arrives at the situation that the origin falls inside a neighborhood of φ_0 with angular amplitude σ.

To show the existence of $\widetilde{n}$, note that the sequence $(2\pi\frac{\omega_0}{\omega}k)_{k\in\mathcal{Z}_=}$, thought of as a sequence of angular coordinates on the circle, corresponds to a family of points which are pairwise distinct since

$$2\pi\frac{\omega_0}{\omega}k_1 = 2\pi\frac{\omega_0}{\omega}k_2 + 2\pi\mu \qquad (2.20.12)$$

with k_2, k_2, μ integers would imply, if $k_1 \neq k_2$, that $\omega_0/\omega = \mu/(k_1 - k_2) =$ (rational number). Then, if $\overline{\varphi}$ is an accumulation point of the above family of points on the circle, there must exist two distinct points in such a sequence closer to $\overline{\varphi}$ than $\sigma/2$; i.e., there exist $k_1, k_2 > 0$ such that

$$\big|(2\pi\frac{\omega_0}{\omega}k_1 - 2\pi\frac{\omega_0}{\omega}k_2) \bmod 2\pi\big| < \sigma, \qquad (2.20.13)$$

and this means that the rotation $R^{k_1-k_2}$ displaces the point 0 on the circle at a point whose angular distance from 0 is ε and $0 < |\varepsilon| < \sigma$ [note that $\varepsilon \neq 0$, by the remark related to Eq. (2.20.12)]. Hence, one can take $\widetilde{n} = k_1 - k_2$ if $k_1 > k_2$ or $\widetilde{n} = k_2 - k_1$ if $k_2 > k_1$. mbe

2.20.1 Exercises and Problems

Problems (1)-(9) are inspired by [26] and they aim at providing tools for studying the remaining problems.

1. Let $r > 0$ be an irrational number represented by its continued fraction

$$r = a_0 + \frac{1}{a_1 + \frac{1}{a_2+\frac{1}{a_3+\ldots}}} \equiv \{a_0, a_1, a_2, \ldots\}$$

defined by setting $[x] = (integral\ part\ of\ x)$ and $a_0 = [r], r_1 = (r - a_0)^{-1}, a_1 = [r_1], r_2 = (r_1 - a_1)^{-1}, a_2 = [r_2]$, *etc.* Show that $a_j > 0$, $\forall\, j > 0$. Compute $a_0, a_1, a_2, \ldots$ for $r =$ golden section $= (\sqrt{5} - 1)/2$ (note that $r = 1/(1 + r)$),or $r = (1 + \sqrt{5})/2$ (note that $r = 1 + 1/r$), or $r = \sqrt{2}$ (note that $r = 1 + 1/(1 + r)$), or $r = \pi$ (recall $\pi = 3.141592653589\ldots$ and using a pocket computer to find empirically (i.e., without error estimates) $a_0, a_1, a_2, \ldots, a_8$, one finds $a_0 = 3, a_1 = 7, a_2 = 15, a_3 = 1, a_= 291, \ldots$).

2. In the context of problem 1 let

$$R_k = a_0 + \frac{1}{a_1 + \frac{1}{a_2+\frac{1}{a_3+\ldots\ \ldots+\frac{1}{a_k}}}} \stackrel{def}{=} \{a_0, a_1, a_2, \ldots, a_k\}$$

Show that $R_{2k} < r < R_{2k+1}$ for all $k \geq 0$.

3. In the context of problems 1 and 2 note that if $\{a_1, \ldots, a_k\} = \frac{p'}{q'}$ then $\{a_0, a_1, \ldots, a_k\} = \frac{a_0p'+q'}{p'}$. Deduce from this that a vector $\mathbf{v}_k = (p_k, q_k) \in \mathcal{Z}^2_+$ such that $R_k = \frac{p_k}{q_k}$ can be taken

$$\mathbf{v}_k = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

4. Deduce from problem 3 that $\mathbf{v}_k = a_k \mathbf{v}_{k-1} + \mathbf{v}_{k-2}$, i.e.

$$p_k = a_k p_{k-1} + p_{k-2}, \qquad k > 1$$
$$q_k = a_k q_{k-1} + q_{k-2}, \qquad k > 1$$

(*Hint:* $\begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a_k \\ 1 \end{pmatrix} = a_k \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} a_{k-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; then eliminate the last matrix in the product of matrices appearing in problem 3).

5. From the recursion relation in problem 4 deduce that

$$q_k p_{k-1} - p_k q_{k-1} = -(q_{k-1} p_{k-2} - p_{k-1} q_{k-2}) = (-1)^k, \qquad k \geq 2$$
$$q_k p_{k-2} - p_k q_{k-2} = a_k (q_{k-1} p_{k-2} - p_{k-1} q_{k-2}) = (-1)^{k-1} a_k, \qquad k \geq 2$$

so that

$$\frac{p_{k-1}}{q_{k-1}} - \frac{p_k}{q_k} = \frac{(-1)^k}{q_k q_{k-1}}, \qquad \frac{p_{k-12}}{q_{k-2}} - \frac{p_k}{q_k} = \frac{(-1)^{k-1}}{q_k q_{k-2}} a_k$$

(*Hint:* Multiply the first equation in the recursive formula in problem 4 by q_{k-1} and the second by p_{k-1} and subtract, *etc.*)

6. From problem (5) deduce that

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} \ldots < r < \ldots \frac{p_3}{q_3} < \frac{p_1}{q_1}, \qquad \text{and } \left| r - \frac{p_k}{q_k} \right| < \frac{1}{q_k q_{k+1}}$$

7. Show that $q_k \geq 2^{(k-1)/2},\ k \geq 0$ and $p_k \geq 2^{(k-2)/2},\ k \geq 1$. (*Hint:* Note that $a_k \geq 1$ for all $k \geq 1$ and use the recursive relation in problem 4 and $p_j, q_j \geq 1$.)

8. Show that

$$\frac{1}{q_k(q_k + q_{k+1})} < \left| r - \frac{p_k}{q_k} \right| < \frac{1}{q_k q_{k+1}}$$

(*Hint:* If $\frac{a}{b} < \frac{c}{d}$ then $\frac{a + s\,c}{b + s\,d}$ increases with s for $s \geq 0$, while if $\frac{a}{b} > \frac{c}{d}$ it decreases. Hence if k is even $\frac{p_{k-2} + s\,p_{k-1}}{q_{k-2} + s\,q_{k-1}}$ increases with s and for $s = a_k$ it becomes $\frac{p_k}{q_k}$ which is such that $\frac{p_k}{q_k} < r < \frac{p_{k-1}}{q_{k-1}}$. Therefore

$$\frac{p_{k-2}}{q_{k-2}} \leq \frac{p_{k-2} + p_{k-1}}{q_{k-2} + q_{k-1}} < r$$

hence

$$\left| r - \frac{p_{k-2}}{q_{k-2}} \right| > \left| \frac{p_{k-2} + p_{k-1}}{q_{k-2} + q_{k-1}} - \frac{p_{k-2}}{q_{k-2}} \right| \equiv \frac{1}{q_{k-2}(q_{k-2} + q_{k-1})})$$

9. Show that the numbers p_n, q_n are relatively prime for all n. (*Hint:* this is obvious for p_0, q_0; suppose it is true for $p_k, q_k,\ k = 0, 1, 2, \ldots, n$ and remark that if p', q' are the last convergents of the continued fraction $[a_1, a_2, \ldots, a_n]$ then they are by the inductive assumption relatively prime and $p_n = a_0 p' + q'$, $q_n = p'$; hence if j divided q_n and p_n, it would divide p' and q', against the assumption on p', q').

The following definition will be used below: a rational number p/q is a best approximation for r if for any pair p', q' with $q' < q$ it is $|q'r - p'| > |qr - p|$.

10. Let p, q be positive integers and assume r irrational. Let j be odd, $\alpha = p/q$, $\alpha_j = p_j/q_j$, and suppose that $\alpha_{j-1} < \alpha < \alpha_{j+1}$; then $q > q_j$. (*Hint:* $\alpha_{j-1} > \alpha > \alpha_{j+1} > r > \alpha_j$ so that $(q_j q_{j-1})^{-1} > |\alpha_{j-1} - r| > |\alpha_{j-1} - \alpha| = |p_{j-1}q - q_{j-1}p|/qq_{j-1} \geq 1/qq_{j-1}$, because $|p_{j-1}q - q_{j-1}p| \geq 1$). State and check the analogous result for j odd, showing that the two results can be summarized by saying that if p/q is between two convergents of orders $j-1$ and $j+1$ then $q > q_j$.

11. In the context of problem (10) show that if α is not a convergent and $\alpha_{j-1} < \alpha < \alpha_{j+1}$ then $q_j\, |r - \alpha_j| < q\, |r - \alpha|$; a similar result holds for j even. (*Hint:* $q|\alpha - r| > q|\alpha - \alpha_{j+1}| = q|pq_{j+1} - qp_{j+1}|/qq_{j+1} \geq 1/q_{j+1} \geq q_j|\alpha_j - r|$).

12. Show that problems (9),(10,(11) imply that if p/q us an approximation to r such that $|q'r - p'| > |qr - p|$ for all $q' < q$ then $q = q_j$, $p = p_j$ for some j. In other words every best approximant is a convergent.

13. Show that if r is irrational every convergent is a best approximant. (*Hint:* if not it must be that that for some n there exists $q < q_n$ with $|rq - p| < |rq_n - p_n| = \varepsilon_n$; let $\bar{p}, \bar{q}$ minimize the expression $|q'r - p'|$ for $q' < q_n$; if $\bar{\varepsilon}$ is the minimum value, it is $\bar{\varepsilon} < \varepsilon_n$; hence $\bar{p}/\bar{q}$ is a best approximation: so that $\bar{p} = p_s, \bar{q} = q_s$ for some $s < q_n$ and $1/(q_s + q_{(s+1)}) \leq |q_s r - p_s| \leq |q_n r - p_n| < 1/q_{n+1}$, i.e. $q_s + q_{s+1} > q_{n+1}$ which contradicts $q_{n+1} = a_{n+1}q_n + q_{n-1}$).

14. A necessary and sufficient condition in order that a rational approximation to an irrational number be a best approximation is that it is a convergent of the continued fraction of r. (*Hint:* just a summary of Problems (9)-(13)).

15. Show that if $q_{n-1} < q < q_n$ then $|qr - p| > |q_{n-1}r - p_{n-1}|$. (*Hint:* if not and if $\bar{\varepsilon} = \min |qr - p|$ over $q_{n-1} < q < q_n$ and over p is reached at some $\bar{q}, \bar{p}$ then $\bar{p}/\bar{q}$ would be a best approximation). Show that this can be interpreted as saying that the graph of the function $\eta(q) = \min_p |qr - p|$ is above that of the function $\eta_0(q) = \varepsilon_n = |q_n r - p_n|$ for $q_n <\leq q < q_{n+1}$.

16. Suppose n even and think the interval $[0, 1]$ as a circle of radius $1/2\pi$: the point $q_n r \mod 1$ can be represented as a point displaced by ε_n to the right of 0, while $q_{n-1}r$ can be viewed as a point to the left of 0 by ε_{n-1}. Show that the points qr with $q_n < q < q_{n+1}$ are not in the interval $[0, \varepsilon_{n-1}]$ unless q/q_n is an integer $\leq a_{n+1}$. Furthermore show that the point $a_{n+1}q_n r$ is closer than ε_n to ε_{n-1}, and that the *next* position closest to 0 occurs when the *rotation* by $(q_n a_{n+1} + q_{n-1})r \equiv q_{n+1}r$ is considered and it is to the left of 0 and at a distance ε_{n+1} from it. Show that this provides a natural interpretation of the meaning of the numbers a_j in the continued fraction of r regarded as a rotation of the circle $[0, 1]$, as well as a geometric interpretation of the relation $a_{n+1}q_n + q_{n-1} = q_{n+1}$.

17. Show that the function $\varepsilon(T) =$ *maximum gap between points of the form* $nr \mod 1, n = 1, 0, \ldots, T$ depends on T as

$$\begin{aligned}
q_n \leq T < q_n + q_{n-1} \ &\Rightarrow\ \varepsilon(T) = \varepsilon_{n-1}\\
q_n + q_{n-1} \leq T < 2q_n + q_{n-1} \ &\Rightarrow\ \varepsilon(T) = \varepsilon_{n-1} - \varepsilon_n\\
\ldots\ &\Rightarrow\ \ldots\\
(a_{n+1} - 1)q_n \leq T < a_{n+1}q_n + q_{n-1} \equiv q_{n+1} \ &\Rightarrow\ \varepsilon(T) = \varepsilon_{n-1} - (a_{n+1} - 1)\varepsilon_n
\end{aligned}$$

and apply this to draw the diagram of $\varepsilon(T)$ and its inverse $T(\varepsilon)$ for the *golden number*, i.e. the number with $a_j \equiv 1$. Plot $-\log \varepsilon(T)$ in terms of $\log T$. (*Hint:* this is simply another interpretation of problem (16)).

18. Show that if the entries a_j of the irrational number r are uniformly bounded by N then the growth of q_n is bounded by an exponential (and one can estimate q_n by a constant

times $[(N+(N^2+4)^{1/2})/2]^n$). Vice-versa an exponential bound can hold if and only if the entries of the continued fraction are uniformly bounded. (*Hint:* it is bounded by the q_n of the number with continued fraction with entries all equal to N).

19. Show that if the inequality: $|q_n r - p_n| > 1/Cq_n$ holds for all n and for a suitable C then q_n cannot grow faster than exponential. (*Hint:* Problem (8) implies the inequality $1/Cq_n < 1/q_{n+1}$.)

20. Show that if a number has a continued fraction with entries which eventually are periodically repeated, then it is a number verifying a quadratic equation, i.e. it is a *quadratic irrational.* Vice-versa it can also be shown that all quadratic irrationals have a continued fraction with entries eventually periodic repeated. (See problems (21),(22) below).

21. Suppose that for some integers a, b, c it is $ar^2 + br + c = 0$. Remark that the argument in Problem (3) shows that the number $r_n = [a_n, a_{n+1}, \ldots]$ verifies $r = (p_{n-1}r_n + p_{n-2})/(q_{n-1}r_n + q_{n-2})$. Substituting the latter expression in the equation for r one finds that r_n verifies an equation like $A_n r_n^2 + B_n r_n + C_n = 0$. Check, by direct calculation of A_n, B_n, C_n that:

$$A_n = a\,p_{n-1}^2 + b\,p_{n-1}q_{n-1} + c\,q_{n-1}^2$$
$$C_n = A_{n-1}$$
$$B_n^2 - 4A_n\,C_n = b^2 - 4a\,c$$

Show that $|A_n|, |B_n|, |C_n|$ are uniformly bounded by $H = 2(2|a|r + |b| + |a|)) + |b|$.(*Hint:* it suffices to find a bound for $|A_n|$. Write $A_n = q_{n-1}^2(a(p_{n-1}/q_{n-1})^2 + b(p_{n-1}/q_{n-1}) + c)$ and use that $|r - p_{n-1}/q_{n-1}| < 1/q_{n-1}^2$ and $ar^2 + br + c = 0$).

22. Show that a quadratic irrational has an eventually periodic continued fraction because, as a consequence of the results of the previous problem, the numbers r_n can only take finitely many values. Show that, if H is the constant introduced in problem (21), the period length can be bounded by $2(2H+1)^3$ and that the periodic part has to start from the j-th entry with $j \le 2(2H+1)^3$.

23. Let $\omega = \{a_0, a_1, \ldots, a_k\}$, $a_i \ge 1, i > 0$, be a rational number and let $\boldsymbol{\omega} = (\omega, 1)$. Consider the periodic motion on $\mathcal{T}^2$ given by $\boldsymbol{\alpha}_0 + t\,\boldsymbol{\omega}$. Estimate (from below) the maximum distance that a point can have from the trajectory of $\boldsymbol{\alpha}_0$.

24. Determine the region Ω densely covered by the data-space trajectory of the motion $\ddot{x} + x = 3\cos\omega t$, $\dot{x}(0) = x(0) = 0$, when ω is irrational.

25. For $\omega =$ golden section (see Problem (1)) estimate the minimum time τ necessary for the trajectory of the oscillator in Problem (24) to cover Ω so that any point in Ω has a distance from the trajectory $t \to (\dot{x}(t), x(t))$, $t \in [0, T]$, not exceeding $\sigma = 2\pi/2^4$.

26. Same as Problem (25) for $\omega = \sqrt{2}$ and for $\omega = \pi$; (use for π an empirically computed continued fraction; see Problem (1)).

27. Let $\widetilde{\omega} = \{a_0, a_1, \ldots, a_k\}$, $a_i \ge 1$, $i > 0$, be a rational number. In terms of k, estimate the maximum distance of a point in Ω from the trajectory of the oscillator in Problem (24) with $\widetilde{\omega}$ replacing ω.

2.21 Quasi-Periodic Functions. Multi Periodic Functions. Tori and the Multidimensional Fourier Theorem

The considerations of 2.20 suggest the following definition.

11 Definition. *A function* $f \in C^\infty(\mathcal{R})$ *is "quasi-periodic with pulsations* $\omega_1, \ldots, \omega_d$*" if there exists a* $\varphi \in C^\infty(\mathcal{R}^d)$ *such that*

$$\varphi(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_d) = \varphi(\alpha_1, \ldots, \alpha_i + 2\pi, \ldots, \alpha_d), \tag{2.21.1}$$

$\boldsymbol{\alpha} \in \mathcal{R}^d$, $i = 1, 2, \ldots, d$, *and*

$$f(t) = \varphi(\omega_1 t, \ldots, \omega_d t), \qquad t \in \mathcal{R}. \tag{2.21.2}$$

The numbers $T_1 = 2\pi/\omega_1, \ldots, 2\pi/\omega_d$*, are called the "periods" of* f*, while* $\nu_1 = T_1^{-1}, \ldots, \nu_d = T_d^{-1}$ *are the "frequencies" of* f*.*

Observations.
(1) Therefore the motion of a harmonic oscillator with pulsation ω_0 forced by a periodic force with pulsation ω is, in absence of resonances, a quasi-periodic function with pulsations ω_0 and ω [see Eq. (2.16.6) and §2.20].
(2) The above definition of a quasi-periodic function is more restrictive than the one sometimes found in mathematical literature: it is, however, sufficiently general for our purposes.
(3) It is useful to note that given f, there exist several choices of $d, \omega_1, \ldots, \omega_d$ and φ allowing us to represent f as in Eq. (2.21.2). A trivial example is provided by the consideration of a function $\varphi \in C^\infty(\mathcal{R})$, periodic with period 2π, and of the functions of $\xi \in \mathcal{R}$ or of $(\xi_1, \xi_2) \in \mathcal{R}^2$ defined as $\psi(\xi)$ or $\widetilde{\psi}(\xi_1, \xi_2) = \varphi(2\xi_1 + 3\xi_2)$ which, via the formulae

$$f(t) = \psi(\frac{\omega}{2}t) \equiv \varphi(\omega t), \tag{2.21.3}$$

$$f(t) = \widetilde{\psi}(\frac{\omega}{4}t, \frac{\omega}{6}t) \equiv \varphi(\omega t), \tag{2.21.4}$$

allow a representation of f as a quasi-periodic function with angular velocities $\omega/2$ or ω or $(\omega/4, \omega/6)$.
(4) The pulsations (or "angular velocities") in Definition 11 need not necessarily all be positive: some may be zero or negative.

The functions φ used to introduce the notion of quasi-periodic function are remarkable in themselves, and it is convenient to set up the following definition.

12 Definition. *Given* $L_1, \ldots, L_d > 0$*, consider the pavement of* $\mathcal{R}^d$ *whose tesserae are the parallelepiped* $[0, L_1] \times [0, L_2] \times \ldots \times [0, L_d]$ *and the parallelepipeds obtained by translating it by* $(n_1 L_1, \ldots, n_d L_d)$*,* $n_1, \ldots n_d$ *integers. Two points* $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathcal{R}^d$ *will be declared equivalent if they are "equally located" in the pavements tesserae, i.e., if there are* d *integers* $n_1, \ldots, n_d$ *such that* $\xi_i - \eta_i = n_i L_i$*,* $i = 1, \ldots, d$*. Then* $\mathcal{T}^d(L_1, \ldots, L_d)$ *will denote the set of the equivalence classes thus obtained and a "distance" will be defined as*

$$d(\{\boldsymbol{\xi}\},\{\boldsymbol{\eta}\}) \stackrel{def}{=} \min_{\substack{\boldsymbol{\xi}'\in\{\boldsymbol{\xi}\}\\ \boldsymbol{\eta}'\in\{\boldsymbol{\eta}\}}} |\boldsymbol{\xi}'-\boldsymbol{\eta}'| \tag{2.21.5}$$

if $\{\boldsymbol{\xi}\},\{\boldsymbol{\eta}\} \in \mathcal{T}^d(L_1,\ldots,L_d)$ *and* $\{\boldsymbol{\xi}\}$ *denotes the equivalence class containing* $\boldsymbol{\xi}$*. The set* $\mathcal{T}^d(L_1,\ldots,L_d)$*, regarded as a metric space with the distance function defined by Eq. (2.21.5) ("distance on the torus"), will be called a "d-dimensional torus" with sides* $L_1,\ldots,L_d$*. If* $L_1 = L_2,\ldots = L_d = 2\pi$*, this torus will be denoted* $\mathcal{T}^d$*, simply, and called "standard torus".*

Observations.

(1) The above definition, in spite of its apparent complexity, is simple and can be informally summarized by saying that the torus $T^d(L_1,\ldots,L_d)$ is obtained by "identifying the opposite sides" of the parallelepiped $[0,L_1]\times\ldots\times[0,L_d]$ of $\mathcal{R}^d$. For this reason it is customary to describe points of $\mathcal{T}^d(L_1,\ldots,L_d)$ through the Cartesian coordinates in $\mathcal{R}^d$ of one of the corresponding representatives without explicit mention of the equivalence relation: when $L_1 = \ldots = L_d = 2\pi$, such coordinates are called the "natural angular coordinates" or "flat coordinates" on $\mathcal{T}^d$. In general, the distance [Eq. (2.21.5)] is called the distance between $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ on the torus $\mathcal{T}^d(L_1,\ldots,L_d)$.
(2) Clearly $\mathcal{T}^d$ can be regarded as the product of d unit circles. If $(\varphi_1,\ldots,\varphi_d)$ are the natural angular coordinates of $\boldsymbol{\varphi}\in\mathcal{T}^d$, a natural one-to-one continuous mapping of $\mathcal{T}^d$ into $S\times S\times\ldots\times S$, where $S =$ (unit circle in the complex plane), is

$$\boldsymbol{\varphi} = (\varphi_1,\ldots,\varphi_d)\longleftrightarrow \mathbf{z} = (z_1,\ldots,z_d) = (e^{i\varphi_1},\ldots,e^{i\varphi_d}) \tag{2.21.6}$$

and the distance (2.21.5) turns out to be equivalent to the distance on $S\times S\times\ldots\times S$ as a subset of $\mathcal{C}^d$. Therefore, the d-dimensional torus $\mathcal{T}^d$ can be regarded as a subset of the d-dimensional complex space $\mathcal{C}^d$. This representation is more intrinsic since it does not involve coordinates defined $\mod 2\pi$. It will turn out to be a deep and very useful representation (see Chapter V, §5.10-5.12).

13 Definition. $C^\infty(\mathcal{T}^d(L_1,\ldots,L_d))$ *is, by definition, the set of the functions* f *defined on* $\mathcal{T}^d(L_1,\ldots,L_d)$ *such that setting (notations of Definition 12)*

$$\widetilde{f}(\boldsymbol{\xi}) = f(\{\boldsymbol{\xi}\}), \qquad \forall\,\boldsymbol{\xi}\in\mathcal{R}^d \tag{2.21.7}$$

the function $\widetilde{f}$ *is in* $C^\infty(\mathcal{R}^d)$*. The set of functions on* $\mathcal{R}^d$ *having the form of Eq. (2.21.7) is the set of the "multi periodic functions on* $\mathcal{R}^d$*" with periods* $L_1,\ldots,L_d$*.*
When $\widetilde{f}$ *has the form of Eq. (2.21.7) with* $f\in C^\infty(\mathcal{T}^d(L_1,\ldots,L_d))$*, the same happens for the partial derivatives of* f *since the derivatives of a* C^∞*-multi periodic function are still multi periodic; i.e., given* d *nonnegative integers* $n_1,\ldots,n_d$*, there is* $\varphi_{n_1,\ldots,n_d}\in C^\infty(\mathcal{T}^d d(L_1,\ldots.L_d))$ *such that*

$$\frac{\partial^{n_1+\ldots+n_d}\widetilde{f}}{\partial\xi_1^{n_1},\ldots,\xi_d^{n_d}}(\boldsymbol{\xi})=\varphi_{n_1,\ldots,n_d}(\{\boldsymbol{\xi}\}) \tag{2.21.8}$$

and it is natural to set

$$\frac{\partial^{n_1+\ldots+n_d}f}{\partial\xi_1^{n_1},\ldots,\xi_d^{n_d}}(\boldsymbol{\xi})=\varphi_{n_1,\ldots,n_d}(\{\boldsymbol{\xi}\}) \tag{2.21.9}$$

Depending on the circumstances, it is possible to think or not to think of a C^∞-multi periodic function with periods $L_1,\ldots,L_d$ and its partial derivatives as an element of $C^\infty(\mathcal{T}^d(L_1,\ldots,L_d))$.

Observations.

(1) Another natural definition of $C^\infty(\mathcal{T}^d)$, for $L_1=\ldots=L_d=2\pi$, could be related to observation (2) to Definition 12: one could say that $f\in C^\infty(\mathcal{T}^d)$ if $f(\boldsymbol{\varphi})=F(\mathbf{z})$, where F is a C^∞-function on $\mathcal{C}^d$[14] and $\mathbf{z}$ is given by Eq. (2.21.6). This would in fact be an equivalent definition, as could be shown; see Problems (6)-(10) at the end of this section.

(2) Along the same lines, after Definition 13, one can define the classes $C^\infty(V\times\mathcal{T}^d)$, where V is an open set in $\mathcal{R}^q$, and the derivatives of their elements. One can also define $W\times\mathcal{T}^\ell$-valued functions in $C^\infty(V\times\mathcal{T}^d)$, and their derivatives, as the $C^\infty(V\times\mathcal{T}^d)$ functions with values in $\mathcal{R}^s\times\mathcal{R}^\ell$ whose last ℓ components are thought of as angular coordinates on $\mathcal{T}^\ell$ (for $W\subset\mathcal{R}^s$, $V\subset\mathcal{R}^q$ open sets).

(3) An example of a multi periodic function on $\mathcal{R}^d$ with periods $\frac{2\pi}{\omega_1},\ldots,\frac{2\pi}{\omega_p}$ is the sum of the series

$$f(\xi_1,\ldots,\xi_d)=\sum_{n_1,\ldots,n_d}^{n_j\in\mathcal{Z}}c_{n_1,\ldots,n_d}e^{i(n_1\omega_1\xi_1+\ldots+n_d\omega_d\xi_d)}; \tag{2.21.10}$$

provided the coefficients $c_{n_1,\ldots,n_d}\in\mathcal{C}$ verify ("reality condition")

$$c_{n_1,\ldots,n_d}=\overline{c}_{-n_1,\ldots,-n_d} \tag{2.21.11}$$

and if, $\forall\, s=0,1,\ldots$, there exists $\gamma_s>0$ such that ("regularity condition")

$$(1+|n_1|)^s\ldots(1+|n_d|)^s|c_{n_1,\ldots,n_d}|\le\gamma_s \tag{2.21.12}$$

The partial derivatives of f, in the sense of Definition 12, can be computed by series differentiation, as a result of Eq. (2.21.12).

It is important to realize that, vice versa, Eqs. (2.21.10), (2.21.11), and (2.21.12) provide the most general example. This is essentially the content of the following proposition ("multidimensional Fourier theorem").

[14] A C^∞-function on $\mathcal{C}^d$ is a function on $\mathcal{C}^d$ which is C^∞ in the real and imaginary parts of the coordinates.

28 Proposition. *Let f is a C^∞-multi periodic function on $\mathcal{R}^d$ with periods $L_1, \ldots, L_d > 0$, then it is possible to represent f by formula (2.21.10) with coefficients $c_{n_1,\ldots,n_d}$ verifying Eqs. (2.21.11) and (2.21.12) and given by*

$$c_{n_1,\ldots,n_d} = \int_0^{L_1} \frac{d\xi_1}{L_1} \int_0^{L_2} \frac{d\xi_2}{L_2} \cdots \int_0^{L_d} \frac{d\xi_d}{L_d} \cdot e^{-i\,(\omega_1 n_1 \xi_1 + \ldots + \omega_d n_d \xi_d)}\, f(\xi_1, \ldots, \xi_d), \tag{2.21.13}$$

where $\omega_j = 2\pi/L_j$, $j = 1, \ldots, d$.

Observations.

(1) If $d = 1$, this proposition coincides with the Fourier development theorem for periodic functions (see Proposition 19).
(2) Since $\widetilde{f}(\xi_1, \ldots, \xi_d) = f(\xi_1/\omega_1, \ldots, \xi_d/\omega_d)$ is multi periodic with periods 2π, it will suffice to prove the above proposition when $\omega_1 = \ldots = \omega_d = 1$.

Proof (Case $\omega_1 = \ldots = \omega_d = 1$). The proof can be developed by induction. For $d = 1$ it holds (see Proposition 19, §2.12); hence assume its validity for $d = 1, 2, \ldots, k$ and consider the case $d = k + 1$.

Let $f \in C^\infty(\mathcal{T}^{k+l})$ and contemplate the function $\psi_{\xi_{k+1}}$, parameterized by $\xi_{k+1} \in \mathcal{R}$ and defined on $\mathcal{R}^k$:

$$\psi_{\xi_{k+1}}(\xi_1, \ldots, \xi_k) = f(\xi_1, \ldots, \xi_k, \xi_{k+1}), \tag{2.21.14}$$

which, $\forall\, \xi_{k+1} \in \mathcal{R}$, is a C^∞-2π-multi periodic function on $\mathcal{R}^k$. The inductive hypothesis implies

$$f(\xi_1, \ldots, \xi_k, \xi_{k+1}) = \sum_{(n_1,\ldots,n_k) \in \mathcal{Z}^k} \widehat{\psi}_{n_1,\ldots,n_k}(\xi_{k+1}) e^{i\,(n_1\xi_1 + \ldots + n_k \xi_k)} \tag{2.21.15}$$

with

$$\widehat{\psi}_{n_1,\ldots,n_k}(\xi_{k+1}) = \int_0^{2\pi} \frac{d\xi_1 \ldots d\xi_k}{(2\pi)^k} f(\xi_1, \ldots, \xi_k, \xi_{k+1}) e^{-i\,(n_1\xi_1 + \ldots + n_k \xi_k)}) \tag{2.21.16}$$

On the other hand, Eq. (2.21.16) immediately implies that $\widehat{\psi}_{n_1,\ldots,n_k}(\xi_{k+1})$ is a C^∞-function, periodic with period 2π, of ξ_{k+1}, for all choices of $(n_1, \ldots, n_k)$ $\in \mathcal{Z}^k$. Via the Fourier theorem for $d = 1$ it follows

$$\widehat{\psi}_{n_1,\ldots,n_k}(\xi_{k+1}) = \sum_{n_{k+1} \in \mathcal{Z}} e^{in_{k+1}\xi_{k+1}} \int_0^{2\pi} \widehat{\psi}_{n_1,\ldots,n_k}(\xi') e^{-in_{k+1}\xi'} \frac{d\xi'}{2\pi} \tag{2.21.17}$$

i.e., using Eq. (2.21.13) as definition of $c_{n_1,\ldots,n_{k+1}}$ and inserting Eq. (2.21.16) in the right-hand side of Eq. (2.21.17), we find

$$\widehat{\psi}_{n_1,\dots,n_k}(\xi_{k+1}) = \sum_{n_{k+1}\in\mathcal{Z}} c_{n_1,\dots,n_{k+1}} e^{in_{k+1}\xi_{k+1}} \tag{2.21.18}$$

Substituting Eq. (2.21.18) into Eq. (2.21.15), one obtains Eq. (2.21.10), provided Eq. (2.21.12) holds (which implies that the series on n_{k+1}, and on $n_1,\dots,n_k$ can be unconditionally summed and interchanged since they are absolutely convergent).

It is therefore necessary to check Eq. (2.21.12) in order to complete the proof. In fact, Eq. (2.21.11) follows from Eq. (2.21.13) which has now become, temporarily, a definition of $c_{n_1,\dots,n_{k+1}}$. One can proceed as in the analogous situation met in the one-dimensional case: one integrates Eq. (2.21.13) by parts. By integrating σ times with respect to ξ_j by parts, we find, if $n_j \neq 0$:

$$c_{n_1,\dots,n_{k+1}} = \frac{1}{(in_j)^\sigma} \int_0^{2\pi} \frac{d\xi_1 \dots d\xi_{k+1}}{(2\pi)^{k+1}} \cdot \frac{\partial^\sigma f(\xi_1,\dots,\xi_k,\xi_{k+1})}{\partial \xi_j^\sigma} e^{-i\sum_r n_r\xi_r} \tag{2.21.19}$$

and, if $F'_\sigma = \max_{\xi,j} |\frac{\partial^\sigma f}{\partial \xi_j^\sigma}(\boldsymbol{\xi})|$, this yields

$$|c_{n_1,\dots,n_{k+1}}| \leq \frac{F'_\sigma}{|n_j|^\sigma} \tag{2.21.20}$$

For $n_j = 0$, from Eq. (2.21.13), and $\forall\, n_1,\dots,n_{k+1}$, (zero or not), $c_{n_1,\dots,n_{k+1}}$ is bounded by the maximum $F'_{\mathbf{0}}$ of $|f|$, Eq. (2.21.20) implies the existence of some $F_\sigma > 0$ such that for $\forall\, j = 1,\dots,k+1$, $\forall\, \sigma \in \mathcal{Z}_+$, $\forall\, (n_1,\dots,n_{k+1}) \in \mathcal{Z}^{k+1}$:

$$|c_{n_1,\dots,n_{k+1}}| \leq \frac{F_\sigma}{(1+|n_j|)^\sigma} \tag{2.21.21}$$

(take, for instance, $F_0 = F'_{\mathbf{0}} + F'_\sigma$). Hence, multiplying Eq. (2.21.21) on j as j varies between 1 and $k+1$ and then taking the $(k+1)$-th root, side by side, of the result, it is

$$|c_{n_1,\dots,n_{k+1}}| \leq \frac{F_\varphi}{[(1+|n_1|)\dots(1+|n_{k+1}|)]^{-\sigma/(k+1)}}, \tag{2.21.22}$$

implying Eq. (2.21.12) by the arbitrariness of $\sigma > 0$. mbe

It is useful to explicitly state the following obvious corollary of Proposition 28 and Definition 11.

29 Corollary. *If f is a C^∞-quasi-periodic function with pulsations $\omega_1,\dots,\omega_d > 0$, then it admits a representation of the type*

$$f(t) = \sum_{\mathbf{n}\in\mathcal{Z}^d} c_{\mathbf{n}} e^{i\,\mathbf{n}\cdot\boldsymbol{\omega}\,t}, \tag{2.21.23}$$

where $\boldsymbol{\omega} = (\omega_1,\dots,\omega_d)$, $\mathbf{n} = (n_1,\dots,n_d)$, and the constants $c_{n_1,\dots,n_d}$ verify Eqs. (2.21.11) and (2.21.12).

It is remarkable that in some cases, given $\omega_1, \ldots, \omega_d$, the representation Eq. (2.21.23) is unique.

30 Proposition. *Let $f \in C^\infty(\mathcal{R})$ be quasi periodic with pulsations $\omega_1, \ldots, \omega_d > 0$. If the pulsations are rationally independent,*[15] *the coefficients of the representation (2.21.23) are given by*

$$c_{\mathbf{n}} = \lim_{t \to +\infty} t^{-1} \int_0^t e^{-i\mathbf{n}\cdot\boldsymbol{\omega}\,\tau} f(\tau) d\tau \qquad (2.21.24)$$

and, therefore, the representation (2.21.23) is unique, given $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)$.

PROOF. Taking into account the decay properties of $c_{\mathbf{n}}$ as $n \to \infty$ expressed by Eq. (2.21.12), the integral in Eq. (2.21.24) can be computed via the series in Eq. (2.21.23):

$$t^{-1} \int_0^t e^{-i\mathbf{n}\cdot\boldsymbol{\omega}\,\tau} f(\tau) d\tau = \sum_{\mathbf{m} \in \mathcal{Z}^d} c_{\mathbf{m}} t^{-1} \int_0^t e^{-i(\mathbf{n}-\mathbf{m})\cdot\boldsymbol{\omega}\tau} d\tau. \qquad (2.21.25)$$

The right-hand integral divided by t has a modulus ≤ 1 (as an average of a function with modulus 1). Therefore, Eq. (2.21.12) shows that the series in Eq. (2.21.25) is a series uniformly convergent with respect to t and that the limit of Eq. (2.21.24) can be computed in Eq. (2.21.25) by interchanging it with the series. If $\mathbf{n} \neq \mathbf{m}$, the integral in Eq. (2.21.25) is

$$\frac{1}{t} \frac{e^{-i(\mathbf{n}-\mathbf{m})\cdot\boldsymbol{\omega}t} - 1}{-i(\mathbf{n}-\mathbf{m})} \xrightarrow[t\to\infty]{} 0 \qquad (2.21.26)$$

because $(\mathbf{n}-\mathbf{m})\cdot\boldsymbol{\omega} \neq 0$ by the rational independence assumption on $\boldsymbol{\omega}$.

Hence all the terms in Eq. (2.21.25) with $\neq\mathbf{m}$ do not contribute to the limit, as $t \to +\infty$, of Eq. (2.21.15) itself. The term with $\mathbf{n} = \mathbf{m}$, on the other hand, only contributes $c_{\mathbf{n}}$; hence, the proposition is proved. mbe

For the sake of completeness, we also wonder about what can be said in the other cases when $\omega_1, \ldots, \omega_d$ are rationally dependent. As an example, the following proposition will be discussed,

31 Proposition. *Let $f \in C^\infty(\mathcal{R})$ be quasi-periodic with rationally dependent pulsations $\omega_1, \ldots, \omega_d > 0$. There exist $p < d$ and p rationally independent numbers $\widetilde{o}_1, \ldots, \widetilde{\omega}_p$ and a multi periodic function $\widetilde{\varphi} \in C^\infty(\mathcal{T}^p)$ such that*

$$f(t) = \widetilde{\varphi}(\widetilde{\omega}_1 t, \ldots, \widetilde{\omega}_p), \quad \forall\, t \in \mathcal{R}. \qquad (2.21.27)$$

[15] A family $\Omega = (\omega_1, \omega_2, \ldots)$ of real numbers is said to consist of rationally independent numbers when every finite subset $(\omega_{j_1}, \ldots, \omega_{j_p})$ has the property that the relation $\sum_{k=1}^p n_k \omega_{j_k} = 0$, with $n_1, \ldots, n_p$ integers, implies $n_1 = \ldots = n_p = 0$.

Observation. Therefore, if $\omega_1, \ldots, \omega_d$ are rationally dependent, it is possible to reduce the complexity of the representation Eq. (2.21.23) by reducing the dimension of the multiple series appearing in it.

PROOF. Consider all the subsets of $\omega_1, \ldots, \omega_d$ built with rationally independent numbers and let $(\overline{\omega}_1, \ldots, \overline{\omega}_p)$ be a maximal one among them (i.e., such that $(\overline{\omega}_1, \ldots, \overline{\omega}_p, \omega')$ is not built with rationally independent numbers no matter which $\omega' \in (\omega_1, \ldots, \omega_d)$ is chosen).

Without loss of generality, suppose $\omega_1 = \overline{\omega}_1, \ldots, \omega_p = \overline{\omega}_p$: then for every $j = p+1, \ldots, d$, there are p rational numbers $\Gamma_1^{(j)}, \ldots, \Gamma_p^{(j)}$, all with the same denominator N, as it can and and shall be assumed, such that

$$\omega_j = \sum_{k=1}^{p} \Gamma_k^{(j)} \omega_k, \qquad j = p+1, \ldots, d. \tag{2.21.28}$$

Hence, setting $\Gamma_k^{(j)} = m_k^{(j)}/N$, $m_k^{(j)}$ integer, $j = p+1, \ldots, d$, $k = 1, \ldots, p$, and $\widetilde{\omega}_j \omega_j/N$, we see that

$$\omega_j = \sum_{k=1}^{p} m_k^{(j)} \widetilde{\omega}_k, \qquad j = p+1, \ldots, d. \tag{2.21.29}$$

defining $m_k^{(j)} = N\delta_{jk}$ for $j \le p$. Now make use of Proposition 29 to get

$$\begin{aligned}
f(t) &= \sum_{n_1,\ldots,n_d} c_{n_1,\ldots,n_d} e^{i \sum n_k \omega_k \, t} \\
&= \sum_{n_1,\ldots,n_d} c_{n_1,\ldots,n_d} e^{i \sum_{h=1}^{d} n_h (\sum_{k=1}^{p} m_k^{(h)} \widetilde{\omega}_k) \, t} \\
&= \sum_{n_1,\ldots,n_d} c_{n_1,\ldots,n_d} e^{i \sum_{k=1}^{p} \widetilde{\omega}_k (\sum_{h=1}^{d} m_k^{(h)} n_h) \, t} \\
&= \sum_{q_1,\ldots,q_p} \Big(\sum_{\substack{n_1,\ldots,n_d \\ \sum_h m_k^{(h)} n_h = q_k}} c_{n_1,\ldots,n_d} \Big) e^{i \sum_{k=1}^{p} \widetilde{\omega}_k q_k \, t},
\end{aligned} \tag{2.21.30}$$

Therefore, we set

$$\widetilde{c}_{q_1,\ldots,q_p} = \sum_{\substack{n_1,\ldots,n_d \\ \sum_h m_k^{(h)} n_h = q_k}} c_{n_1,\ldots,n_d}, \tag{2.21.31}$$

and, from Eq. (2.21.11), t is seen that $c_{q_1,\ldots,q_p} = \overline{c}_{-q_1,\ldots,-q_p}$. Furthermore, since $|q_j| \le M(\sum_{k=1}^{d} |n_k|)$ with $M = \max_{k,j} |m_j^{(k)}| \ge 1$, we see that

$$(1+|q_1|)^s \dots (1+|q_p|)^s |\widetilde{c}_{q_1,\dots,q_p}| \le M^{ps} \sum_{\substack{n_1,\dots,n_d \\ \sum_h m_k^{(h)} n_h = q_k}} (1+|q_1|)^{sp} \dots (1+|q_p|)^{sp} |c_{n_1,\dots,n_d}| \tag{2.21.32}$$

The series on the right-hand side of Eq. (2.21.32) can be bounded, with the help of Eq. (2.21.12), as

$$\sum_{\substack{n_1,\dots,n_d \\ \sum_h m_k^{(h)} n_h = q_k}} \frac{(1+|q_1|)^{sp} \dots (1+|q_p|)^{sp} \,\gamma_{sp+2}}{(1+|q_1|)^{sp+2} \dots (1+|q_p|)^{sp+2}} \le \Big(\sum_{n=-\infty}^{+\infty} \frac{2}{(1+|n|)^2} \Big)^d \gamma_{sp+2} \tag{2.21.33}$$

Hence, Eqs. (2.21.32) and (2.21.33) mean that the constants c'' verify an inequality like Eq. (2.21.12) (with p instead of d) and the proposition is now proved since, by Eq. (2.21.30), we can define

$$\widetilde{\varphi}(\xi_1,\dots,\xi_p) = \sum_{q_1,\dots,q_p} \widetilde{c}_{q_1,\dots,q_p} e^{i\sum_{kj=1}^{p} \widetilde{\omega}_j \, q_j \, \xi_j} \tag{2.21.34}$$

mbe

2.21.1 Exercises and Problems

1. Compute the Fourier coefficients $\widetilde{f}_{0,0}, \widetilde{f}_{0,1}, \widetilde{f}_{1,0}$ of the function $f(\xi_1,\xi_2) = 1 - \frac{1}{4}(\cos\xi_1 + \cos\xi_2)^{-1}$ with an approximation of 50%.

2. Same as Problem 1 for $f(\xi_1,\xi_2) = 1 - \log(\cos\xi_1 + \cos\xi_2)$.

3. Show that if $f(\xi_1,\xi_2) = \sum_{k=0}^{\infty} 4^{-k} C_k (\cos\xi_1 + \cos\xi_2)^k$ with $|C_k| < D$, there exist $C > 0, \varepsilon > 0$ such that $|\widetilde{f}_{n_1,n_2}| \le C e^{-\varepsilon(|n_1|+|n_2|)}$. Estimate C and ε in terms of D.

4. If ω_1/ω_2 is irrational, show that, for $\forall\, \varphi \in C^\infty(\mathcal{T}^2)$, the closure of the set of the values taken as $t \in \mathcal{R}_+$, by $f(t) = \varphi(\omega_1 t, \omega_2 t)$ coincides with $\varphi(\mathcal{T}^2) = \varphi$-image of $\mathcal{T}^2$. (*Hint:*: See Proposition 27.)

5.* Same as Problem 4 when $\varphi \in C^\infty(\mathcal{T}^d)$, $f(t) = \varphi(\omega_1 t, \dots, \omega_d t)$ and $\omega_1, \dots, \omega_d$ are rationally independent.

6. On the complex plane $\mathcal{C}/\{0\}$, define the function $I(z) = e^{i\varphi}$ if $z = \varrho e^{i\varphi} \ne 0$, $\varrho > 0, \varphi \in \mathcal{R}$. Show that I is a C^∞ function of $\mathcal{R}\mathrm{e}\, z = x$ and $\mathcal{I}\mathrm{m}\, z = y$.

7. In the context of Problem 6, show that

$$\Big|\frac{\partial^k I(z^n)}{\partial x^h \partial y^{k-h}}\Big| \le |n|^k C_k$$

For a suitably chosen C_k, for all z such that $\frac{1}{2} \le |z| \le 2$.

8. If $f \in C^\infty(\mathcal{R})$ and f is 2π-periodic and if $\widetilde{f}_n$ are the Fourier coefficients of f, consider the function of $z = x + iy, x, y \in \mathcal{R}$ defined for $\frac{1}{2} < |z| < 2$ and by

$$F(z) = \sum_{n=-\infty}^{+\infty} \widetilde{f}_n \, I(z^n).$$

Using Problem 7, show that F, as a function of x, y, is C^∞ in the region $\frac{1}{2} < |z| < 2$ and on the unit circle coincides with $f(\varphi) = F(e^{i\varphi})$.

9. Using Problem 8, show the validity of the equivalence claimed in observation (1) to Definition 13, p. 102, in the case $d = 1$.

10. Same as Problem 9 in the case $d > 1$. (*Hint:*: If $f \in C^\infty(\mathcal{T}^d)$ and if $\widetilde{f}_{n_1,\ldots,n_d}$ are its Fourier coefficients, let $\mathbf{z} = (z_1, \ldots, z_d) \in \mathcal{C}^d$ and

$$F(\mathbf{z}) = \sum_{n_1,\ldots,n_d \in \mathcal{Z}^d} \widetilde{f}_{n_1,\ldots,n_d} I(z_1^{n_1}) \cdots I(z_d^{n_d});$$

then show that F is a C^∞ function of $x_i = \mathcal{R}\mathrm{e}\, z_i$ and $y_i = \mathcal{I}\mathrm{m}\, z_i$, $i = 1, \ldots, d$, in a neighborhood of the torus $S \times \ldots \times S$ where $S = \{\text{unit circle in } \mathcal{C}\}$ identified with $\mathcal{T}^d$ via Eq. (2.21.6).)

2.22 Observables and Their Time Averages

Observables and time averages play an important role in qualitative as well as quantitative developments in Mechanics. It is therefore useful to look more closely at them. For the purpose, consider an autonomous differential equation

$$m\,\ddot{x} = f(\dot{x}, x), \tag{2.22.1}$$

where $(\eta, \xi) \to f(\eta, \xi)$ is in $C^\infty(\mathcal{R}^2)$ and $m > 0$. Suppose, also, that Eq. (2.22.1) is normal. According to Definition 7, we shall denote by $(S_t)_{t \in \mathcal{R}_+}$, the flow which solves Eq. (2.22.1); i.e., if $(\eta, \xi) \in \mathcal{R}^2$, the function

$$t \to (\dot{x}(t), x(t)) = S_t(\eta, \xi), \qquad t \in \mathcal{R}_+ \tag{2.22.2}$$

will be such that $t \to x(t)$, $t \in \mathcal{R}_+$, is a solution of Eq. (2.22.1) with initial datum (η, ξ). Recall, also, that the map defined on $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ and with values in $\mathcal{R} \times \mathcal{R}$ is a C^∞ map and

$$S_{t+t'} = S_t\, S_{t'}, \qquad \forall\, t, t' \in \mathcal{R}_+; \tag{2.22.3}$$

see Corollary 9. In the above context, introduce the following concepts.

14 Definition. *The set of C^∞ functions on $\mathcal{R}^2$, thought of as the space of the initial data of Eq. (2.22.1), will be called the set of instantaneous "observables" for the point mass described by Eq. (2.22.1).*

If $t \to S_t(\eta, \xi)$, $t \geq 0$, is a motion of Eq. (2.22.1) and if F is an observable, let the "value" of the observable at time $t \in \mathcal{R}_+$ on the motion with initial datum (η, ξ) be

$$F(\dot{x}(t), x(t)) = F(S_t(\eta, \xi)). \tag{2.22.4}$$

The function $t \to F(S_t(\eta, \xi))$, $t \in \mathcal{R}_+$, *is the "history" of the observable* F *on the motion with data* (η, ξ).

Observations.
(1) The motivation for the above terminology is clear. What perhaps needs a few words of comment is why one defines an observable as a function of velocity and position only, see Eq. (2.22.4), rather than, more generally, as a function of acceleration and higher derivatives as well.
Actually, such a definition would not be more general since, via Eq. (2.22.1) and by what it has been observed in §2.4, it is possible to compute all the derivatives of $t \to x(t)$ successive to the first by repeatedly differentiating both sides of Eq. (2.22.1), once $x(t)$ and $\dot{x}(t)$ are known.
(2) Therefore, the observables correspond to physical entities measurable by observing velocity and position of the point mass at a given instant: they are a mathematical model of such entities.

Given an observable F and a motion $t \to S_t(\eta, \xi)$, $t \in \mathcal{R}_+$, one can raise several questions about the observations of F at various times. As an example, the notion of average value of an observable on a given motion will be discussed below.

It is important to remember and to stress that, concerning the notion of the average value of an observable, it is possible to repeat what has already been said about the notion of the stability of equilibrium. It makes no sense to provide an absolute definition of average value of an observable as time elapses. In fact, it is possible to give several meanings to this concept, each corresponding to different needs that may naturally emerge in applications.

Here and in the following sections, only a few interesting examples of definition of time averages will be examined, leaving it to the reader to imagine applications in which a particular definition may appear as a relevant one. The reader should also try to imagine other definitions and the corresponding situations to which they could naturally apply: the methods explained below could then be used to elucidate their properties.

15 Definition. *Let* $F \in C^\infty(\mathcal{R}^2)$ *be an observable for the motions described by (2.22.1) and let* $T > 0$. *We define the continuous average value of* F *on the motion with initial datum* $(\eta, \xi) \in \mathcal{R}^2$ *and on the time interval* $[0, T]$ *as*

$$M_T(F; \eta, \xi) = \frac{1}{T} \int_0^T F(S_t(\eta, \xi))\, dt \tag{2.22.5}$$

The "continuous average value" of F *on the motion with initial datum* (η, ξ) *will be, whenever it exists, the limit*

$$\overline{F}(\eta, \xi) = \lim_{T \to +\infty} M_T(F; \eta, \xi). \tag{2.22.6}$$

Similarly, one could define the average value with observation step $a > 0$:

16 Definition. *If $F \in C^\infty(\mathcal{R}^2)$ is an observable for the motions described by Eq. (2.22.1) and if N is a positive integer, the discrete average value with time step a of F on the motion with initial datum (η, ξ) and relative to N observations, is defined as*

$$M_N^{(a)}(F;\eta,\xi) = \frac{1}{N}\sum_{j=0}^{N-1} F(S_{ja}(\eta,\xi)). \tag{2.22.7}$$

The "discrete average value" with step a of F on the motion $t \to S_t(\eta,\xi)$, $t \in \mathcal{R}_+$, is defined by the limit, whenever it exists,

$$\overline{F}^{(a)}(\eta,\xi) = \lim_{N\to+\infty} M_N^{(a)}(F;\eta,\xi). \tag{2.22.8}$$

Why should one refrain from considering a more general notion?

17 Definition. *If $\varphi \in C^\infty(\mathcal{R}_+)$ and if $T > 0, N \in Z_+, a > 0$ let*

$$\begin{aligned} \mathcal{M}_T(\varphi) =& \frac{1}{T}\int_0^T \varphi(t)dt, \qquad && \overline{\varphi} = \lim_{T\to+\infty} \mathcal{M}_T(\varphi) \\ \mathcal{M}_N^{(a)}(\varphi) =& \frac{1}{N}\sum_{j=0}^{N-1}\varphi(ja), && \overline{\varphi}^{(a)} = \lim_{N\to+\infty}\mathcal{M}_N^{(a)}(\varphi) \end{aligned} \tag{2.22.9}$$

whenever the limits exist. The quantities defined in Eq. (2.22.9) will be called the "continuous average of φ on $[0,T]$", the "continuous average of φ", the "discrete average of φ on N observations with time step a", and the "discrete average of φ with time step a".

Observations.
(1) If φ is constant, $\overline{\varphi} = \overline{\varphi}^{(a)} \equiv \varphi$.
(2) If $\lambda = \lim_{t\to+\infty}\varphi(t)$ exists, then $\overline{\varphi} = \overline{\varphi}^{(a)} = \lambda$: in fact, note that $\mathcal{M}_T(\varphi) - \lambda = \mathcal{M}_T(\varphi - \lambda)$ and if T_ε is such that, $\forall\, t \ge T_\varepsilon$, $|\varphi(t) - \lambda| < \varepsilon$, one has

$$\mathcal{M}_T(\varphi - \lambda) = \frac{1}{T}\int_0^{T_\varepsilon}(\varphi(\tau) - \lambda)\, d\tau + \frac{1}{T}\int_{T_\varepsilon}^{T}(\varphi(\tau) - \lambda)\, d\tau \tag{2.22.10}$$

and the first term in the right-hand side of Eq. (2.22.10) goes to zero as $T \to \infty$, while the second is bounded by $T^{-1}(T - T_\varepsilon)\varepsilon < \varepsilon$. Hence $\lim_{T\to\infty}\mathcal{M}_T(\varphi - \lambda) = 0$ by the arbitrariness of ε, and $\overline{\varphi} = \lambda$. Similarly, one checks that $\overline{\varphi}^{(a)} = \lambda$.
(3) If $\varphi \in C^\infty(\mathcal{R})$ is periodic with period $T_\varphi > 0$,

$$\lim_{T\to\infty}\mathcal{M}_T(\varphi) = \overline{\varphi} = \frac{1}{T_\varphi}\int_0^{T_\varphi}\varphi(\tau)d\tau. \tag{2.22.11}$$

In fact, if $T = nT_\varphi, +\theta$ with n integer and $\theta \in [0, T_\varphi]$, it follows that $T \to \infty \longleftrightarrow n \to \infty$ and

$$\mathcal{M}_T(\varphi) = \frac{1}{nT_\varphi + \eta}\big(n\int_0^{T_\varphi} \varphi(\tau)\,d\tau + \int_0^\theta \varphi(\tau)\,d\tau\big) \tag{2.22.12}$$

implying Eq. (2.22.11).

(4) If $\varphi \in C^\infty(\mathcal{R})$ is periodic with period $T_\varphi > 0$ and if $a > 0$ is such that $T_\varphi/a = p/q$ with p and q relatively prime integers (i.e., if T_φ/a is rational), it follows that

$$\overline{\varphi}^{(a)} = \sum_{m=-\infty}^{+\infty} \widehat{\varphi}_{mp}, \qquad \text{and} \qquad \overline{\varphi}^{(a)} = \frac{1}{p}\sum_{j=0}^{p-1} \varphi(ja) \tag{2.22.13}$$

where $\widehat{\varphi}_n$ are the harmonics of φ relative to the period T_φ. The first relation in Eq. (2.22.13) can be proved as in (3) above. To prove the second, note that

$$\begin{aligned}\mathcal{M}_N^{(a)}(\varphi) =& \frac{1}{N}\sum_{j=0}^{N-1} \varphi(ja) = \frac{1}{N}\sum_{j=0}^{N-1}\sum_{n\in\mathcal{Z}} \widehat{\varphi}_n e^{\frac{2\pi\, i}{T_\varphi} ja} \\ =& \sum_{n\in\mathcal{Z}} \widehat{\varphi}_n \big(\frac{1}{N}\sum_{j=0}^{N-1} e^{\frac{2\pi\, i}{T_\varphi} ja}\big)\end{aligned} \tag{2.22.14}$$

and the term in brackets has modulus < 1 (as an average of numbers with modulus not exceeding 1). Hence, the series in Eq. (2.22.14) is uniformly convergent in N and the limit as $N \to \infty$ can be taken term by term. As already remarked (see Eq. (2.21.26)) if $e^{2\pi\, i\, na/T_\varphi} \neq 1$, one finds

$$\frac{1}{N}\sum_{j=0}^{N-1} e^{2\pi\, i\, n \frac{a}{T_\varphi}\, j} = \frac{1}{N}\frac{e^{2\pi\, i\, n\frac{a}{T_\varphi}\, N} - 1}{e^{2\pi\, i\, n\frac{a}{T_\varphi}} - 1} \xrightarrow[N\to+\infty]{} 0, \tag{2.22.15}$$

while if $e^{2\pi\, i\, na/T_\varphi} \equiv 1$, i.e., if na/T_φ is an integer (i.e., $n = mp$ for some $m \in \mathcal{Z}$), the sum (2.22.15) is clearly 1, identically, $\forall\, N$. Hence, by taking the limit as $N \to \infty$ in Eq. (2.22.14), Eq. (2.22.13) follows.

(5) If $\varphi \in C^\infty(\mathcal{R})$ is T_φ-periodic, $T_\varphi > 0$, and if T_φ/a is irrational, then

$$\overline{\varphi}^{(a)} = \overline{\varphi} = \frac{1}{T_\varphi}\int_0^{T_\varphi} \varphi(\tau)\,d\tau. \tag{2.22.16}$$

This is true because, in the present case, in the series (2.22.14), all the terms tend to zero except the one with $n = 0$ (as $\exp(2\pi i n a/T_\varphi) \neq 1$, $\forall\, n \neq 0$ [see, also, Eq. (2.22.15)]).

(6) If $\varphi \in C^\infty(\mathcal{R})$ is periodic with period $T_\varphi > 0$, let $a > 0$ vary so that T_φ/a is rational, but if $T_\varphi/a = p/q$, with p and q relatively prime integers, then $p \to \infty$.[16] Then it follows from Eq. (2.22.13) and from the decay 112 properties

[16] The number p measures the number of times it is necessary to repeat a to reach a multiple of T_φ, i.e. it measures the "commensurability" of T_φ with respect to a.

of the Fourier coefficients of of the Fourier coefficients of φ that $\overline{\varphi}^{(a)} \to \overline{\varphi}$. Hence, the "less T_φ is commensurable with a", the closer the discrete average $\overline{\varphi}^{(a)}$ is to the continuous average $\overline{\varphi}$.

The following proposition is a consequence of the above remarks and an example of questions related to the corresponding definitions,

32 Proposition. *Let $V \in C^\infty(\mathcal{R})$ be bounded below and consider the motions associated with Eq. (2.22.17):*

$$m\,\ddot{x}(t) = -\frac{\partial V}{\partial \xi}(x(t)), \qquad t \in \mathcal{R}_+. \tag{2.22.17}$$

If F is an observable "with bounded support" (i.e., if $F(\eta, \xi) \equiv 0$ when $|\eta| + |\xi|$ is large enough), every initial datum $(\eta, \xi) \in \mathcal{R}^2$ gives rise to a motion on which both the continuous and the discrete averages with step $a > 0$ are defined.
If $\lim_{\xi \to \pm\infty} V(\xi) = +\infty$, every observable (whether with bounded support or not) has well-defined average values, continuous and discrete. In this case the continuous and discrete averages with step $a > 0$ coincide on all motions, with the possible exception of the periodic motions with period commensurable with a.

PROOF. From Proposition 11, p. 37, it follows that the motions described by Eq. (2.22.17) either approach infinity or tend toward a well-defined limit (i.e., $\lim_{t \to +\infty} S_t(\eta, \xi) = (0, \xi_0)$) or are periodic.

In the first two cases, the above proposition follows from observation 2 to Definition 17, while in the third case, it follows from Observations 3 and 4. The assumption on the support of F is needed to deal with the case when $S_t(\eta, \xi) \to \infty$: this case cannot occur, according to the law of conservation of energy, when V diverges at infinity; hence, in this case, no restriction on F is necessary. mbe

2.22.1 Exercises and Problems

1. Compute the continuous average along the motions $\ddot{x} + x = 0$, $x(0) = 0$, and $\dot{x}(0) = 1$ of the kinetic energy and of the squared elongation (Le., of the observables $f(\eta, \xi) = \frac{1}{2}\eta^2$ or $g(\xi, \eta) = \xi^2$).

2. Compute the difference between the continuous average of the kinetic energy and that of the potential energy in the oscillations of $m\ddot{x} = -kx$ with energy E. Compute their values as functions of E.

3. Compute the discrete average of the kinetic energy for the motion $\ddot{x} + x = 0$, $x(0) = 0, \dot{x}(0) = 1$ for $a = 2\pi, 4\pi, \frac{\pi}{2}, 1, 2, \frac{17}{13}$.

4. Same as Problem 1 for the motion $\ddot{z} + \sin x = 0$, $x(0) = 0$, $\dot{x}(0) = \frac{1}{2}$ with 60% accuracy.

5. Same as Problem 3 for the motion in Problem 4 with 60% accuracy.

6. Same as Problems 4 and 5 with 1% accuracy (using a computer).

7. Same as Problem 1 for the motion in Problem 4. Estimate the accuracy needed in the computations to see a difference between the linear-oscillator and pendulum results.

8. Compute the average value of the elongation, and of the square of the elongation, in the motion $\ddot{x} + x = cos\pi t$, $x(0) = 0, \dot{x}(0) = 0$ in the continuous case and in the discrete case with step $a = \pi, \frac{17}{13}, \sqrt{2}$.

9. Show that if $\varphi, \psi \in C^\infty(\mathcal{R})$ and $\lim_{t\to+\infty} |\varphi(t) - \psi(t)| = 0$ and if φ has an average value of any type, then ψ has the same average value.

10. Apply Problem 9 to calculate the continuous average of the squared elongation in the motion of the oscillator $\ddot{x} + \dot{x} + x = \cos t$, $x(0) = 0, \dot{x}(0) = 0$. How does this average change by changing the initial datum? (*Answer:* It does not change.)

11. Define work per unit time of a force f on a point with velocity v the quantity $v \cdot f$: see p.144 for the general definition. Arbitrarily choose a definition of average and estimate the average work done by the friction force ("dissipation per unit time", i.e., average of the observable. $w(\eta, \xi) = -\eta^2$) in the motions of the oscillator in Problem 10.

12. In the context of Problem 11, compare the average work per unit time done by the friction force and that done by the forcing force. Interpret the value of their difference.

13. Compute, in general, the continuous average value of the work done by the forcing force and by the friction force in the motions of the oscillators $m\ddot{x} + \lambda\dot{x} + kx = f(t)$ with $m, \lambda, k > 0$ and $f(t) = F\cos\omega t$, $F, \omega \in \mathcal{R}$. Also compute the continuous average value of the potential or kinetic energy.

14.* Same as Problem 13 but with a generic $2\pi/\omega$-periodic C^∞ forcing force f. Express the results by means of the harmonics of f and of the parameters m, λ, k.

15. In the context of Problem 13, find the value of ω to which corresponds maximum average work done by the forcing term ("resonant pulsation").

16.* If $f \in C^\infty\mathcal{R}$ is a quasi-periodic function in the sense of Definition 11, then the average values of f exist both in the continuous and the discrete sense. Find expressions for such quantities and show that if the pulsations of f are $\omega_1, \ldots, \omega_d$ and if $\{\omega_1, \ldots, \omega_d, 2\pi/a\}$ are $(d+1)$ rationally independent numbers, then the discrete average of f with step $a > 0$ and the continuous average of f coincide. (*Hint:*: Use the representation of Eq. (2.21.23) and proceed as in Observation 4, Eq. (2.22.14).)

17. Find an example of a function in $C^\infty(\mathcal{R})$ which does not have a continuous average.

18. Estimate within 60% the average kinetic energy in the motion with energy $E = 10$ of the oscillator $\ddot{x} + x^3 = 0$.

19.* Same as Problem 18 with 1% accuracy (using a computer).

20.* Show that if a potential energy produces periodic motions with period $T(E)$ which, as E varies in $[E_0, E_1]$, is such that $T'(E) > 0$, then the discrete average with step $a = 1$ and the continuous average of an arbitrary observable coincide for a dense set of values of $R \in [E_0, E_1]$, while they do not coincide, in general, on another dense set. The second set is, however, denumerable. (*Hint:*: By the implicit functions theorem, deduce that $T(E)/a$ is irrational for all but countably many values of $E \in [E_0, E_1]$).

21. Show that the same results of Problem 20 hold if $T(E)$ is strictly monotonically increasing in $[E_0, E_1]$. They also hold if $T'(E) = 0$ only finitely many times in $[E_0, E_1]$.

2.23 Time Averages on Sequences of Times known up to Errors. Probability and Stochastic Phenomena

... or mi di' anche:
Questa Fortuna di che tu mi tocche,
Che è, che i ben del mondo ha sì tra branche? [17]

The continuous averages as well as the step-a discrete averages are, as is easily understood, very idealized mathematical notions, even when T or N are $< +\infty$. To be really measured, the continuous averages would demand an infinity of measurements of f, one per each time, and there is no need to underline the degree of abstraction that must be assumed in order to imagine such a sequence of measurements.

Only at first sight are the discrete averages "more concrete" notions. It is in fact unthinkable to be able to perform measurements at time intervals exactly equal to a, because of the unavoidable errors of time measurement.

Obviously, considerations of measurement errors could have been brought up in correspondence with almost every question studied so far or it could be brought up in correspondence with any future question. Arbitrarily, we decide to discuss it now in connection with the analysis of the averages of functions or observables.

The methods and ideas involved in the effort of making precise the notion of error in the time average computations present the greatest interest and are very general: they could be applied to the consideration of errors in the context of other problems, and the reader could try some of these applications by himself.

A very naive schematization of the data accumulation process for calculatinb an average is the following: one measures[18] f, the function that we want to average, at the initial time $\tau_0 \simeq 0$; then we wait a time interval $\tau_1 \simeq a$ and repeat, again, the measurement of f, and subsequently the operation is repeated after waiting times $\tau_2, \tau_3, \ldots$ etc: every τ_i, $i = 1, 2, \ldots$ is approximately equal to a, though not exactly because of the errors made in the measurement of the time intervals. Afterwards, the average of f will be defined as the "average of the results thus obtained". Such an average, instead of being

$$\mathcal{M}_N^{(a)}(f), \qquad \text{will be} \tag{2.23.1}$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} f(\tau_0 + \tau_1 + \ldots + \tau_j) \tag{2.23.2}$$

[17] In basic English:
... now tell me also:
This Fortune of whom you speak
What is she, that the world's goods holds so firmly in her hands?
(Dante, Inferno, Canto VII).

[18] For the sake of simplicity, ability to perform exact measurements of f will be supposed so that the only source of error comes from the measurement of the time intervals.

Time measurement errors will be further idealized by imagining that

$$\tau_0 = \varepsilon_0, \quad \tau_1 = a + \varepsilon_1, \qquad i = 1, 2, \ldots, \tag{2.23.3}$$

and $\varepsilon_j = \pm\varepsilon$ with $\varepsilon > 0$ fixed, $\varepsilon \ll a$, and the sign of ε_j is "randomly chosen".

One can think of a simple mechanism producing a sequence of errors like those in Eq. (2.23.3). Assume that to be also able to perform perfect time measurements, but to proceed deliberately as follows: at the initial time toss a coin and perform a measurement of f at time $\tau_0 = \varepsilon_0$, where $\varepsilon_0 = \varepsilon$ if the result is "heads", while $\varepsilon_0 = -\varepsilon$ if the result is "tails".

At time τ_0 we again toss the coin and perform the measurement of f at time $\tau_0 + \tau_1$, where $\tau_1 = a + \varepsilon_1$ and $\varepsilon_1 = \pm\varepsilon$ according to the result,[19] etc.

One can debate at length on which would be the best mathematical model allowing a satisfactory translation into mathematically clear terms of the just-described sequence of operations. The most interesting mathematical scheme is based on the notion of probability.

18 Definition. *Let $\mathcal{E}$ be a finite set of elements which will be called "possible events". On $\mathcal{E}$, let p be a function on $\mathcal{E}$ with $p(e) \geq 0$ such that*

$$\sum_{e\in\mathcal{E}} p(e) = 1. \tag{2.23.4}$$

The pair $(\mathcal{E}, p)$ will be called a "probability distribution" on $\mathcal{E}$. If $A \subset \mathcal{E}$ is a subset of $\mathcal{E}$, we set

$$p(A) = \sum_{e\in A} p(e) \tag{2.23.5}$$

and we say that $p(A)$ is the probability of A with respect to the distribution $(\mathcal{E}, p)$.

The above notion of probability is precise from a mathematical point of view, but its connection with reality is far less evident. A relation between this definition and the empirical world cannot be established on a deductive basis in the same way as it is not possible to establish deductively the relation between solutions to differential equations and motions of point masses.

The theory of a point mass motion, if identified with the theory of a class of differential equations, appears to us as natural only after long practice and experience in comparing the relations between the mathematical model and the corresponding empirical, i.e., experimental, properties of "real" point masses. In this comparison, one refines both the mathematical intuition on the structure of the solutions of some differential equations and the physical intuition about the nature of motion.

Even a superficial knowledge of the theory of differential equations has the consequence that one cannot avoid observing motions, perhaps unconsciously,

[19] In other words, instead of leaving the "coin tossing" to the measurement instruments, we "do it ourselves".

more and more closely to unveil in them those properties which are suggested by their analytical model as solutions of a differential equation.

Similarly, the notion of probability allows the formulation of mathematical models of stochastic (i.e. random) phenomena and the quantitative evaluation of the probability of classes of events, reaching results such as "that class of events has large probability" or "probability $\frac{1}{7}$", etc. In terms of empirical interpretation, the meaning to attribute to such results becomes clearer and more refined while one proceeds in the applications, and this allows us to think of them again in more intuitive terms, more immediately expressible in an empirical language and in empirical prescriptions.

The key to the empirical interpretation of the notion of probability is the following: consider a "stochastic phenomenon" developing "*following the judgement of Her, which is as hidden as a snake in the grass*",[20] which we imagine "reproducible" and whose possible events form a certain set $\mathcal{E}$. To say that a mathematical model for such a phenomenon is given by the probability distribution $(\mathcal{E}, p)$ means to formulate a law (on an empirical basis) stating that the number of times that in "n trials", or "repetitions of the production of the event", the event $e \in \mathcal{E}$[21] will happen about $p(e)\, n$ times, if n is large, and the deviations from this value are very small, $\ll p(e)\, n$, except in "particularly unlucky" situations which can be disregarded "for all practical purposes".

One can wonder about what could be the predictive power of such a law. This power, in fact, is enormous when it is formulated a priori, i.e., without having first measured the occurrence frequencies of every event of $\mathcal{E}$ over a large number of "trials". The laws of dynamics have the same extent of power when they are applied to cases to which they are believed to be applicable, but for which the actual applicability has not been checked a priori and will be checked only a posteriori (think of the microscopic theory of gases, or of the planetary system theory).

Obviously sometimes a formulated law may be wrong, i.e., the distribution $(\mathcal{E}, p)$ may not be a good model of the stochastic phenomenon in the preceding sense. This may happen for two reasons.

First, the phenomenon may be stochastic but the empirical law on the existence of a well-defined frequency of realization of every possible event may not hold, in the limit of a large number of trials. In mechanics an analogous situation would occur in discovering a point mass for which one could find, after a few direct measurements of force and corresponding acceleration, that the two physical entities are not proportional.

Alternatively, it might happen that the probability law $(\mathcal{E}, p)$, assumed as modelling the phenomenon under analysis, foresees occurrence frequencies different from the observed ones: this circumstance would have the analogue,

---

[20] "Seguendo lo giudicio di costei/ che occulto come in erba l'angue" (Dante, Inferno, Canto VII).

[21] $\mathcal{E}$ could be the six faces of a dice and a "try" could be one tossing of the dice (after suitably "shaking" it); and the produced event would be the upper face of the dice after tossing: if the dice is "fair" then $p(e) = \frac{1}{6}$.

in the mechanics of a point mass, of a case where we had "forgotten" to list some force f among the forces acting on the point.

The discussion on the notion of probability and on its empirical interpretation will be stopped here. One could continue it for much longer, at the risk of making the issue and the content of the analysis increasingly nebulous. In fact it is more useful and constructive to illustrate the content of Definition 18 via a few applications to the problems which interest us.

To have at hand a more flexible language, it is convenient to agree on a few more "simple" definitions. First comes the notion of "random variable".

19 Definition. *Let $(\mathcal{E}, p)$ be a probability distribution.*
(i) Any real function f on $\mathcal{E}$ will be called a "random variable".
(ii) If $a_1, a_2, \ldots, a_{n(f)}$, are the pairwise distinct values taken by $f(e)$ as e varies in $\mathcal{E}$, we shall call $E_1, E_2, \ldots, E_{n(f)}$ the corresponding sets of events of $\mathcal{E}$; i.e., for $i = 1, 2, \ldots, n(f)$ the set E_i consists of those elements $e \in \mathcal{E}$ such that $f(e) = a_i$. The sets $(E_1, \ldots, E_{n(f)})$ are pairwise disjoint and their union is $\mathcal{E}$. Therefore, they form a "partition" P_f of $\mathcal{E}$, which will be called "partition of $\mathcal{E}$ associated with f".
(iii) The "probability distribution" of the random variable f is the probability distribution (I_f, P_f), where I_f has as elements the $n(f)$ sets $E_1, \ldots, E_{n(f)}$ and

$$P_f(E_i) = p(E_i) = \sum_{e \in E_i} p(e) \tag{2.23.6}$$

(iv) More generally, if $\mathcal{P}$ is a partition of $\mathcal{E}$ into n sets $(E_1, \ldots, E_n)$, we shall define $(\mathcal{E}_\mathcal{P}, p_\mathcal{P})$ the "probability distribution associated with $\mathcal{P}$" as being the probability distribution in which the elements of $\mathcal{E}_\mathcal{P}$ are the sets constituting the partition $\mathcal{P}$ and, if $E \in \mathcal{P}$,

$$p_\mathcal{P}(E) = p(E) = \sum_{e \in E} p(e) \tag{2.23.7}$$

Observation. The notion of the probability distribution of a random variable is a relevant one when we are only interested in the random event $e \in \mathcal{E}$ via the value $f(e)$. It is in fact clear that we can identify all the events $e \in \mathcal{P}$ giving rise to the same value of $f(e)$ and call "event" such a collection of events.
Suppose, for instance, performing a measurement of a quantity g and that such a measurement is affected by an error which can be thought of as due to N "causes", all independent from each other and each producing an additive error on the value of g which is $\pm\varepsilon$ with equal probability. A complete description of the error is therefore a N-tuple $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)$ of numbers which take the values $\varepsilon_i = \pm\varepsilon$; the hypothesis of independence and equal probability of the various errors will be translated into a model by saying that all the N-tuples $\boldsymbol{\varepsilon}$ are equally probable; i.e., on the space $\mathcal{E}$ of the 2^N sequences

$\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)$, with $\varepsilon_i = \pm\varepsilon$, the probability distribution[22] $p(\boldsymbol{\varepsilon}) = 2^{-N}$ is defined.
Suppose, however, that we are not interested in knowing the details of the individual errors occurrences but just in the total error:

$$f(\boldsymbol{\varepsilon}) = \sum_{i=1}^{N} \varepsilon_i \tag{2.23.8}$$

This is a random variable on $\mathcal{E}$. It can take the values $N\varepsilon, (N-2)\varepsilon, (-N+2)\varepsilon, -N\varepsilon$, and the value $(N-2k)\varepsilon$ is taken on all the sequences $\boldsymbol{\varepsilon}$ containing exactly k minus signs: call E_k the set of all such sequences. Then the set I_f, in this example, consists of $N+1$ elements $E_0, E_1, \ldots, E_N$ and

$$p_f(E_i) = p(E_i) = \sum_{e \in E_i} \frac{1}{2^N} = \frac{1}{2^N}\binom{N}{i} \tag{2.23.9}$$

The probability distribution (I_f, P_f) can be regarded as a model for the total error without explicit reference to the elementary errors ε_i.

The preceding definition provides a method for building new probability distributions, starting from a given probability distribution. It is useful, in this respect, also to give the following definition providing another way of constructing new probability distributions starting from a given one $(\mathcal{E}, p)$, as suggested by the above observation.

20 Definition. *Let $(\mathcal{E}, p)$ be a probability distribution. Let N be a positive integer. We shall denote $(\mathcal{E}, p)^N$ as the probability distribution on $\mathcal{E}^N$ associating with the event $\mathbf{e} = (e_1, \ldots, e_N) \in \mathcal{E}^N$ the probability $p^{(N)}(e)$:*

$$p^{(N)}(e) = p(e_1)p(e_2)\ldots p(e_N). \tag{2.23.10}$$

The distribution $(\mathcal{E}, p)^N$ will be called the "distribution of N events independently extracted with distribution $(\mathcal{E}, p)$".

This series of definitions, necessary to establish a concise and suggestive language for the formulation of some interesting propositions, will be concluded by describing the important notion of a sequence of random variables converging in probability to a constant limit.

21 Definition. *Let $(\mathcal{E}_N, p_N)$, $N = 1, 2, \ldots$, be a sequence of probability distributions and let f_N be a random variable defined on $\mathcal{E}_N$, $N = 1, 2, \ldots$. The sequence $(f_N)_{N=1}^{\infty}$ of random variables is said to "converge in probability" to a limit $\ell \in \mathcal{R}$ as $N \to \infty$, if* [23]

[22] This is a celebrated error model. It was used by Gauss for his mathematical theory of errors, one of the first grandiose applications of probability theory.

[23] We use the convention that $\{\mathbf{e} | \mathbf{e} \in A, f(\mathbf{e}) \in B\}$ means "subset of A consisting in those $\mathbf{e}$'s such that $f(e) \in B$.

$$\lim_{N\to\infty} p_N(\{\mathbf{e}|\mathbf{e} \in \mathcal{E}_N,\ |f(\mathbf{e}) - \ell > \varepsilon\}) = 0 \tag{2.23.11}$$

for all $\varepsilon > 0$.

Let us provide some examples.

33 Proposition. *Let* $(\mathcal{E}, p)$ *be a probability distribution and let* f *be a random variable on* $(\mathcal{E}, p)$. *Define the random variable* f_N *on* $(\mathcal{E}, p)^N$ *as*

$$f_N(\mathbf{e}) = \frac{1}{N} \sum_{i=1}^{N} f(e_i) \tag{2.23.12}$$

if $\mathbf{e} = (e_1, \ldots, e_N) \in \mathcal{E}^N$. *Then the sequence* f_N *converges in probability to* $\overline{f} = \sum_{e\in\mathcal{E}} p(e)\, f(e)$ *as* $N \to \infty$.

Observations.
(1) This proposition ("law of large numbers") tells us that the average value of a sum of N independent random variables is "almost constant" if N is large or, better, that the probability that such an average value differs from a certain constant $\overline{f}$ by more than a given quantity ε approaches 0 as $N \to \infty$ [see Eq. (2.23.12)].
(2) This proposition clarifies why the quantity $\sum_{e\in\mathcal{E}} p(e)f(e)$ is called the "average value" of the random variable f with respect to the probability distribution $(\mathcal{E}, p)$.

The proof of Proposition 33 relies on a very elementary but very important inequality ("the Chebysčev inequality") which underlies many probabilistic estimates.

34 Proposition. *Let* f *be a random variable with respect to the probability distribution* $(\mathcal{E}, p)$. *Define the "k-th moment" of* f *as*

$$\mu_k(f) = \sum_{s\in\mathcal{E}} |f(e)|^k p(e), \qquad k \in \mathcal{Z}_+ \tag{2.23.13}$$

Then for $k \in \mathcal{Z}_+$ *and* $\delta > 0$,

$$P(\{e \,|\, e \in \mathcal{E}, |f(e)| > \delta\}) \le \frac{\mu_k(f)}{\delta^k} \tag{2.23.14}$$

Proof. By Eq. (2.22.13),

$$\mu_k(f) \ge \sum_{\substack{e\in\mathcal{E}\\ |f(e)|>\delta}} |f(e)|p(e) \ge \delta^k \sum_{\substack{e\in\mathcal{E}\\ |f(e)>\delta}} p(e) = \delta^k p(\{e \,|\, e \in \mathcal{E}, |f(e)| > \delta\}). \tag{2.23.15}$$

mbe

Proof of Proposition 33. By applying the Chebysčev inequality to the random variable $f_N - \overline{f}$, one finds

$$p_N(\{\mathbf{e}\,|\,\mathbf{e}\in\mathcal{E}^N, |f_N(\mathbf{e})-\overline{f}|>\delta\}) \le \frac{\mu_2}{\delta^2} \tag{2.23.16}$$

where

$$\begin{aligned}\mu_2 &= \sum_{\mathbf{e}\in\mathcal{E}^N} p_N(\mathbf{e})(f_N(\mathbf{e})-\overline{f})^2 = \sum_{e_1,\ldots,e_N}\Big(\prod_{i=1}^N p(e_i)\Big)\Big(\frac{1}{N}\sum_{j=1}^N (f(e_j)-\overline{f})\Big)^2\\ &= \frac{1}{N^2}\sum_{j,k=1}^N\sum_{e_1,\ldots,e_N}\Big(\prod_{i=1}^N p(e_i)\Big)\cdot(f(e_j)-\overline{f})(f(e_k)-\overline{f})\\ &= \frac{1}{N^2}\sum_{j=1}^N\sum_{e_1,\ldots,e_N}\Big(\prod_{i=1}^N p(e_i)\Big)(f(e_i)-\overline{f})^2\end{aligned} \tag{2.23.17}$$

since all the terms with $j\ne k$ vanish because, via $\sum_e p(e)=1$, it is

$$\sum_{e_1,\ldots,e_N}\Big(\prod_{i=1}^N p(e_i)\Big)(f(e_j)-\overline{f})(f(e_k)-\overline{f}) \tag{2.23.18}$$

$$= \sum_{e_j,e_k} p(e_j)p(e_k)(f(e_j)-\overline{f})(f(e_k)-\overline{f}) = \Big(\sum_e p(e)(f(e)-\overline{f})\Big)^2 = 0$$

by the definition of $\overline{f}=\sum_{e\in\mathcal{E}} p(e)f(e)$, if $j\ne k$.

The last member of Eq. (2.23.17) can be similarly computed yielding

$$\mu_2 = \frac{1}{N^2}\,N\,\Big(\sum_e p(e)(f(e)-\overline{f})^2\Big) = \frac{\sigma^2}{N}, \tag{2.23.19}$$

where $\sigma^2 = \sum_e p(e)(f(e)-\overline{f})^2$ and the proposition is proved as $\mu_2 \xrightarrow[N\to\infty]{} 0$ [see Eq. (2.23.16)]. mbe

Observation. Note that Eqs. (2.23.16) and (2.23.19) show more: they imply that the probability of the event $|f_N(e)-\overline{f}|>\delta_N$ tends to zero as $N\to\infty$. provided the sequence δ_N is such that $N\delta_N^2 \xrightarrow[N\to\infty]{} 0$, i.e. provided δ_N does not go to zero faster or as $N^{-\frac{1}{2}}$.

Also the problem of the determination of the average value of an observable over a sequence of times succeeding each other at time intervals $a\pm\varepsilon$, where the choice of the sign $\pm$ is a random choice in the sense informally discussed at the beginning of this section, can be easily dealt with by the above techniques.

35 Proposition. *Let $f\in C^\infty(\mathcal{R})$ be a periodic function with period $T>0$. Consider the probability distribution $(\mathcal{E}^N,p_N)$ on the space $\mathcal{E}^N$ of the N-tuples $\boldsymbol{\varepsilon}=(e_0,\ldots,e_{N-1})$, $\varepsilon_i=\pm\varepsilon$, $i=0,\ldots,N-1$ where*

$$p_N(\boldsymbol{\varepsilon}) = 2^{-N}, \qquad \forall \boldsymbol{\varepsilon} \in \mathcal{E}^N. \tag{2.23.20}$$

Given $a > \varepsilon$ with a/ε irrational, consider the random variable on $(\mathcal{E}^N, p_N)$

$$\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) = \frac{1}{N} \sum_{j=0}^{N-1} f(ja + \varepsilon_1 + \ldots + \varepsilon_{N-1}). \tag{2.23.21}$$

Then

$$\lim_{N\to\infty} \widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) = \frac{1}{T} \int_0^T f(\tau)\, d\tau = \overline{f} \tag{2.23.22}$$

in probability.

Observations.
(1) The interest of the proposition is that, even if some measurement errors involving the successive timing of the observations are present, the average value of f, computed using the data successively obtained, has a large probability of being close to the "ideal" average value, i.e., to the continuous average, independent of a and ε, provided a/ε is irrational.
(2) The coincidence of the stochastic average with the continuous average depends upon the irrationality of a/ε, but not on the value of T: it is therefore a property of the structure of the measurement (through the parameters a and ε) and does not depend on the characteristic properties of the observable f; unlike in the comparison between the two ideal notions of the average (continuous and discrete with step a) where the rationality of T/a was relevant (see Proposition 32).
(3) With the same methods of proof, the above proposition could be extended to the case when ε; takes more than two values: $\varepsilon_1 = \pm\alpha_1, \ldots, \pm\alpha_k$, and $p(\alpha_j) \equiv p(-\alpha_j)$. In this case, the condition "ε/a irrational" will be replaced by the condition "there is at least one value $\overline{\alpha}$ among the values of α_j such that $\overline{\alpha}/a$ is irrational".
(4) Finally, always with the same technique of proof, one could treat the case ε/a rational, and this would lead one to conclude that $\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon})$ still converges in probability to a well-defined limit expressible in terms of the Fourier transform of f [with a result analogous to Eq. (2.22.13) generally involving T as well; see Problem. 17 at the end of this section]. The difference between this new limit and the continuous average could be measured by "the commensurability of a with respect to ε" [see observation 6 to Definition 17, p. 111 (for an analogous comment) and Problem 18 at the end of this section].
When the error takes more than one value, as in Observation 3 above, this difference depends on the maximum degree of commensurability between a and the values of the various errors. It is sufficient that among the various errors there is one with respect to which a is "little" commensurable to imply that $\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon})$ converges in probability to a value very close to the continuous average of f.
For this reason, it is rare that the stochastic average sensibly deviates from

the continuous average: in the concrete situations, there are always several causes of errors and, correspondingly, α_j can take very many different values. Necessarily, a will not be too commensurable with respect to many of them.

Proof. The proof is basically a check, relying on the Chebysčev inequality. Consider the Fourier representation of f:

$$f(t) = \sum_{n\in\mathcal{Z}} \widehat{f}_n \, e^{\frac{2\pi\, i}{\,} T\, nt}, \tag{2.23.23}$$

and take into account that $\widehat{f}_0 = \overline{f} = T^{-1}\int_0^T f(\tau)d\tau$:

$$\begin{aligned}
\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) - \overline{f} =& \frac{1}{N}\Big(\sum_{j=0}^{N-1} f(ja+\varepsilon_0+\ldots+\varepsilon_j)\Big) - \overline{f} \\
\equiv& \frac{1}{N}\sum_{j=0}^{N-1}(f(ja+\varepsilon_0+\ldots+\varepsilon_j)-\overline{f}) \\
=& \sum_{\substack{n\in\mathcal{Z}\\ n\neq 0}} \widehat{f}_n\Big(\frac{1}{N}\sum_{j=0}^{N-1} e^{\frac{2\pi\, i}{T} n\,(ja+\varepsilon_1+\ldots+\varepsilon_j)}\Big).
\end{aligned} \tag{2.23.24}$$

Hence, to apply the Chebysčev inequality, compute the second moment of $\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) - \overline{f}$ using Eq. (2.23.24):

$$\begin{aligned}
\mu_2(N) =& \sum_{\boldsymbol{\varepsilon}} (\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) - \overline{f})^2 p_N(\boldsymbol{\varepsilon}) = \frac{1}{2^N}\sum_{\boldsymbol{\varepsilon}}(\widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) - \overline{f})^2 \\
& \sum_{\substack{n_1,n_2\in\mathcal{Z}\\ n_1,n_2\neq 0}} \widehat{f}_{n_1}\widehat{f}_{n_2}\Big\{\frac{1}{N^2}\sum_{j_1,j_2}^{0,N-1}\frac{1}{2^N} e^{\frac{2\pi\, i}{T}(j_1 n_1 a + j_2 n_2 a)} \\
& e^{\frac{2\pi\, i}{T}(\varepsilon_0+\ldots+\varepsilon_{j_1})n_1 + \frac{2\pi\, i}{T}(\varepsilon_0+\ldots+\varepsilon_{j_2})n_2}\Big\}.
\end{aligned} \tag{2.23.25}$$

The series over n_1, n_2 is term-by-term bounded by the convergent series $\sum_{n_1,n_2}|\widehat{f}_{n_1}||\widehat{f}_{n_2}|$: in fact, the factor within curly brackets is a sum of $N^2 2^N$ addends each with modulus $1/N^2 2^N$ and, therefore, its modulus does not exceed 1. Hence, the series in Eq. (2.23.25) is uniformly convergent in N and its limit as $N\to\infty$ can be computed under the summation sign (i.e., term by term). It will turn out that all the terms in curly brackets in Eq. (2.23.25) tend to zero as $N\to\infty$; hence, $\mu_2(N) \xrightarrow[N\to\infty]{} 0$ which, by the Chebysčev inequality, will imply Proposition 35.

The contribution to the sum inside the curly brackets in Eq. (2.23.25) coming from the terms with $j_1 = j_2$ involves $N 2^N$ terms with modulus $1/N^2 2^N$. Hence, it tends to zero as $N\to\infty$. Therefore, it will be enough to consider

the terms with $j_1 < j_2$ and to show that their contribution to the sum is also infinitesimal as $N \to \infty$. The terms with $j_1 > j_2$ can be similarly treated.

Suppose $j_2 > j_1$: the contribution to the curly bracket term from such addends is

$$\begin{aligned}&\frac{1}{N^2}\sum_{j_1=0}^{N-2}\sum_{j_2=j_1+1}^{N-1} e^{\frac{2\pi i}{T}(j_1 n_1 a+j_2 n_2 a)}\\ &\cdot\frac{1}{2^N}\Big\{\sum_{\boldsymbol{\varepsilon}} e^{\frac{2\pi i}{T}(\varepsilon_0+\ldots+\varepsilon_{j_1})(n_1+n_2)+\frac{2\pi i}{T}(\varepsilon_{j_1+1}+\ldots+\varepsilon_{j_2})n_2}\Big\},\end{aligned} \tag{2.23.26}$$

which, by successively performing the summations over $\varepsilon_{N-1},\ldots,\varepsilon_0$, becomes

$$\begin{aligned}&\frac{1}{N^2}\sum_{j_1=0}^{N-2}\sum_{j_2=j_1+1}^{N-1} e^{\frac{2\pi i}{T}(j_1 n_1 a+j_2 n_2 a)}(\cos\frac{2\pi}{T}\varepsilon n_2)^{j_2-j_1}(\cos\frac{2\pi}{T}\varepsilon(n_1+n_2))^{j_1}\\ &=\frac{1}{N^2}\sum_{j_1=0}^{N-2}\sum_{j_2=j_1+1}^{N-1}\big(e^{\frac{2\pi i}{T}(n_1+n_2)a}\cos\frac{2\pi}{T}\varepsilon(n_1+n_2)\big)^{j_1}\big(e^{\frac{2\pi i}{T}n_2}\cos\frac{2\pi}{T}\varepsilon n_2\big)^{j_2-j_1}.\end{aligned} \tag{2.23.27}$$

The summation over j_2 can now be performed, noting that if $n_2 \neq 0$,

$$\lambda = e^{\frac{2\pi i}{T}n_2}\cos\frac{2\pi}{T}\varepsilon n_2 \neq 1 \tag{2.23.28}$$

because $|\lambda| =\leq 1$ and if $\lambda = 1$ the number ε/a would have to be rational, regardless of T. The result of the sum in Eq. (2.23.27) over j_2 is then

$$\frac{1}{N^2}\sum_{j_1=0}^{N-2}\big(e^{\frac{2\pi i}{T}(n_1+n_2)a}\cos\frac{2\pi}{T}\varepsilon(n_1+n_2)\big)^{j_1}\lambda\frac{\lambda^{N-j_1-1}-1}{\lambda-1}, \tag{2.23.29}$$

and this sum involves $(N-1)$ addends each with modulus bounded by $N^{-2}\frac{2}{\lambda-1}$. Hence, it tends to zero as $N \to \infty$. mbe

2.23.1 Exercises and Problems

1. Consider the "fair probability" distribution $(\mathcal{E}, p)$ on a set of six events $\mathcal{E} = \{1, 2, \ldots, 6\}$, $p(j) = \frac{1}{6}$ ("perfect dice"). Compute the probability distribution for the following random variables (see Definition 20):

$$f_1(i) = \begin{cases}1 & \text{if } i \text{ is even,}\\ -1 & \text{if } i \text{ is odd,}\end{cases} \qquad f_2(i) = \begin{cases}1 & \text{if } i = 1,2,3,\\ -1 & \text{if } i = 4,5,6..\end{cases}$$

2. Let $\mathcal{E}$ consist of two elements $+1$ and -1 and let $p(\pm 1) = \frac{1}{2}$. Compute, in $(\mathcal{E}, p)^N$, the moment μ_2 of the random variables $f(\boldsymbol{\varepsilon}) = (\varepsilon_1 + \ldots + \varepsilon_N)$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N) \in \mathcal{E}^N$, and $f(\boldsymbol{\varepsilon}) = \sin(\varepsilon_1 + \ldots + \varepsilon_N)$.

3. Same as Problem 2 with $p(+1) = \frac{2}{3}, p(-1) = \frac{1}{3}$.

4. Compute the limit in probability of the random variables $(\varepsilon_1 + \ldots + \varepsilon_N)/N$, as $N \to +\infty$, in $(\mathcal{E}, p)^N$, where $\mathcal{E} = \{-1, +1\}$, $p(-1) = \frac{1}{3}$, $p(+1) = \frac{2}{3}$.

5. Consider the stochastic average with step $a = \sqrt{2}$ with respect to an error distribution with the scheme $\mathcal{E} = \{-\varepsilon, \varepsilon\}, p(\pm\varepsilon) = \frac{1}{2}, \varepsilon = \frac{1}{10}$ for the observable "kinetic energy" on the energy 1 motion of the oscillator $\ddot{x} + x = 0$. Estimate the number of measurements N needed for finding that the average over N observations deviates from the stochastic average (i.e., from the case $N = +\infty$) by 10% at most with a probability of 99%.

6. Same as Problem 5 with error scheme $\mathcal{E} = \{-\varepsilon, 0, \varepsilon\}$, $p(\pm\varepsilon) = \frac{1}{3}$, $p(0) = \frac{1}{3}$, using the observable "potential energy."

7. Same as Problem 5 for the motions of the oscillators $\ddot{x} + .\dot{x} + x = (1 - \cos t)^2$ for the observable "work done per unit time by the forcing force."

8. Interpret $\sum_{i=1}^{N} \log i \equiv \log n!$ as an approximation for the integral between 1 and $n + 1$ of the function $\xi \to \log \xi$ and, using this interpretation, show that

$$0 < \log n! - n(\log n - 1) \le 1 + \frac{1}{n} + \log n, \quad \text{i.e.} \quad 1 \le \frac{n!}{n^n e^{-n}} \le n\, e^{1 + \frac{1}{n}}.$$

9.* Using the "Stirling formula" (see Problem 14):

$$n! = n^n e^{-n} \sqrt{2\pi n} \Big(1 + O(\frac{1}{n})\Big),$$

show that the probability that

$$f_n(\boldsymbol{\varepsilon}) = \frac{\varepsilon_1 + \ldots + \varepsilon_N}{\sqrt{N}} \in [a, b]$$

with respect to the probability distribution $(\mathcal{E}, p)^N$, where $\mathcal{E} = \{-1, +1\}$, $p(\pm 1) = \frac{1}{2}$, converges to

$$\int_a^b e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

as $N \to +\infty$ ("Gauss' theorem"). (*Hint:* Recall Eq. (2.23.9) to see that the probability that $\frac{\varepsilon_1 + \ldots + \varepsilon_N}{\sqrt{N}}$ takes the value $(N - 2k)/\sqrt{N}$, $k = 0, 1, \ldots, N$, is given by $2^{-N}\binom{N}{k}$; then express the factorials in $\binom{N}{k}$ via the Stirling formula, recalling that k must be such that $a \le (N - 2k)/\sqrt{N} \le b$, etc.)

10.* Show that the statement in Problem 9 implies that the sequence of random variables f_N considered there does not converge in probability as $N \to \infty$.

11. Assuming the result in Problem 9, show that the sequence

$$f_N^{(a)}(\boldsymbol{\varepsilon}) = \frac{\varepsilon_1 + \ldots + \varepsilon_N}{N^{\alpha/2}}$$

of random variables with respect to the probability distribution considered in Problem 9 converges to zero in probability if $\alpha > 1$, does not converge if $\alpha = 1$ (see Problem 10), and diverges if $\alpha < 1$ (in the sense that the probability that $|f_N^{(a)}(\boldsymbol{\varepsilon})| < a$ approaches zero, as $N \to \infty$.)

12. Show that the probability p_N that the random variable $f_N(\boldsymbol{\varepsilon})$ introduced in Problem 9 is positive approaches $\frac{1}{2}$ as $N \to +\infty$ (*Hint:* Distinguish N even and N odd.)

13. In the context of Problems 9 and 12, estimate how fast $|p_N - \frac{1}{2}| \to 0$ as $N \to \infty$.

14.* Prove Stirling's formula with a constant Γ instead of 2π, leaving aside the determination of Γ, refining the argument in Problem 8. (*Hint:*

$$\begin{aligned}\log n!1 =& \sum_{i=2}^{n} \log i = \sum_{i=2}^{n} \int_{i-1}^{i} \log i\, dx = \sum_{i=2}^{n} \int_{i-1}^{i} \log(x-(x-i))\, dx\\ =& \sum_{i=2}^{n} \int_{i-1}^{i} \Big[\log x + \log(1-\frac{x-i}{x})\Big]\, dx = \int_1^n \log x\, dx\\ &+\sum_{i=2}^{n} \int_{i-1}^{i} \Big[\log(1-\frac{x-i}{x}) + \frac{x-i}{x}\Big] dx + \sum_{i=2}^{n} \int_{i-1}^{i} \frac{-(x-i)}{x}\, dx\\ =& n(\log n - 1) + \sum_{i=2}^{n} \gamma_i + \sum_{i=2}^{n} -(1 - i\log\frac{i}{i-1}).\end{aligned}$$

where γ_i denotes the second integral in the intermediate step. Then $|\gamma_i| \le \mathrm{const}\, i^{-2}$ and $-1 + i\log\frac{i}{i-1} = \frac{1}{2i} + \frac{1}{3i^2} + \ldots$ so that

$$\log n! = n(\log n - 1) + \frac{1}{2}\sum_{i=2}^{n}\frac{1}{i} + \sum_{i=2}^{n}\widetilde{\gamma}_i$$

with $\widetilde{\gamma}_i \le \mathrm{const}\, i^{-2}$. Since $\sum_{i=2}^{n}\frac{1}{i} = \log n - \widetilde{C} + O(\frac{1}{n})$ with $\widetilde{C}$ suitably chosen (see next exercise), it follows that

$$\log n! = n(\log n - 1) + \log\sqrt{n} - \widetilde{C} - \sum_{i=2}^{n}\widetilde{\gamma}_i + O(\frac{1}{n});$$

so if $\Gamma = \exp(\widetilde{C} + \sum_{i=2}^{n}\widetilde{\gamma}_i)$, it follows that $n! = n^n e^{-n}\sqrt{n}\,\Gamma\,(1 + O(\frac{1}{n}))$.)

15. Show that $\sum_{i=1}^{n}\frac{1}{i} = \log n - C + O(\frac{1}{n})$, where C is a suitable constant ("Euler-Mascheroni constant") (*Hint:*

$$\begin{aligned}\sum_{i=1}^{n}\frac{1}{i} =& \sum_{i=1}^{n}\int_i^{i+1}\frac{1}{i}\, dx = \sum_{i=1}^{n}\int_i^{i+1}\frac{1}{x+(i-x)}\, dx = \sum_{i=1}^{n}\int_i^{i+1}\frac{1}{1+\frac{(i-x)}{x}}\frac{dx}{x}\\ =& \sum_{i=1}^{n}\int_i^{i+1}\frac{dx}{x}\Big[\frac{1}{1+\frac{i-x}{x}} - 1 + \frac{i-x}{x}\Big] + \sum_{i=1}^{n}\int_i^{i+1}\frac{dx}{x}(1-\frac{i-x}{x}) = \sum_{i=1}^{n}\overline{\gamma}_i + \log(n+1)\end{aligned}$$

and show that $|\overline{\gamma}_i| \le \mathrm{const} i^{-2}$.

16.* Complete the derivation of the Stirling formula begun in Problem 14 by showing that $\Gamma = \sqrt{2\pi}$. (*Hint:* Use Problem 9, with Γ instead of 2π, which says that the random variables $f_N\boldsymbol{\varepsilon})$ lie in $[-A, A]$ with a probability converging to $\int_{-A}^{A} e^{-\frac{x^2}{2}}\, dx/\sqrt{\Gamma}$ (if one does not suppose $\Gamma = \sqrt{2\pi}$ yet). Then, by estimating the factorials in $\binom{N}{k}2^{-N}$ by using the Stirling formula with Γ instead of $\sqrt{2\pi}$, see Problem 14, show that

$$\sum_{|N-2k|/\sqrt{N}>A}\binom{N}{k}2^{-N} \xrightarrow[N\to+\infty]{} 0$$

uniformly in N: this implies that $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\, dx/\sqrt{\Gamma} = 1$; hence, $\Gamma = \sqrt{\pi}$. The estimate on the $\sum 2^{-N}\binom{N}{k}$ is quite delicate and should be decomposed into two estimates: for instance the first for $|\frac{k}{N} - \frac{1}{2}| \in [\frac{A}{\sqrt{N}}, \frac{1}{10}]$ and the second for $|\frac{k}{N} - \frac{1}{2}| \in [\frac{1}{10}, \frac{1}{10}]$.)

17. Show that if a/ε is not assumed to be irrational, Eq. (2.23.22) becomes $\lim_{N\to\infty} \widetilde{\mathcal{M}}_N(\boldsymbol{\varepsilon}) = \sum_n^* \widehat{f}_n = \widetilde{f}$, where $\widehat{f}_n$ are the harmonics of f and $\sum_n^*$ is a sum running over the n's such that $\Big(\exp(\frac{2\pi i}{T} na)\Big) \cos \frac{2\pi}{T} n\varepsilon = 1$

18. Deduce from Problem 17 that the limit in Eq. (2.23.22) coincides in probability with the continuous average not only if a/ε is irrational, but also if T is irrational with respect to either ε or a. Also if $\varepsilon/a = p/q$, with p and q relatively prime integers, and if a is varied so that $p \to \infty, a \to \overline{a} > 0$, then $\widetilde{f} \to \overline{f}$.

2.24 Extremal Properties of Conservative Motion: Action and Variational Principle

Since the construction of the entire universe is absolutely perfect and is due to a Creator with infinite knowledge, nothing exists in the world which does not exhibit some property of maximum or minimum. Therefore, there cannot be any doubt whatsoever about the possibility that all the effects are determined by their final aims with the help of the maxima method, in the same way in which they are also determined by the initial causes.

Equilibrium positions of a point mass on a line are identified with the points where the potential energy is stationary. Thinking of equilibrium as a particular form of motion, one can ask whether the other possible motions of a point mass, developing under the action of a conservative force with potential energy V, can be characterized by similar stationarity properties. This analysis will also be useful as a first illustration of the content of the above quoted proposition of Euler. A deeper analysis will be the object of Chapter 3.

Consider a point with mass $m > 0$ moving in the time interval $[t_1, t_2]$ from the position ξ_1 to the position ξ_2: such a motion is a C^∞ function $t \to x(t)$, $t \in [t_1, t_2]$, such that $x(t_1) = \xi_1, x(t_2) = \xi_2$.

Let $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ be the set of all C^∞ motions $t \to x(t)$, $t \in [t_1, t_2]$, such that $x(t_1) = \xi_1$, $x(t_2) = \xi_2$. If $V \in C^\infty(\mathcal{R})$ is a given function bounded from below, it makes sense to consider the motions of the point taking place under the influence of the force generated by the potential energy V. Such motions are a very restricted class in $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ possibly empty.

The inquiry subject will be whether there is a real-valued function A defined on $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ which takes a minimum value or, at least, is stationary on the motions which, *under the influence of the force with potential energy* V go from ξ_1, to ξ_2 as t goes from t_1, to t_2 .

The meaning of this question has to be clarified by a preliminary discussion on the meaning of "extremality" of a function defined on a set of motions, i.e., on a set of other functions. Attention will focus on special functions defined on $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$: those having the form

$$A(\mathbf{x}) = \int_{t_1}^{t_2} \mathcal{L}(\dot{x}(t), x(t), t)\, dt \tag{2.24.1}$$

where $\mathcal{L} \in C^\infty(\mathcal{R}^3)$ associates (η, ξ, t) with $\mathcal{L}(\eta, \xi, t)$.

Eq. (2.24.1) associates a real number with every $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$. This number is called the "action of the motion $\mathbf{x}$ with respect to the Lagrangian function $\mathcal{L}$.[24]. The notion of "stationarity" or "extremality" of A is very natural in terms of the related notion of "varied motions".

22 Definition. *Given $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ and a real function $(t, \varepsilon) \to y(t, \varepsilon)$ in $C^\infty([t_1, t_2] \times (-1, 1))$ such that*

$$(i) \quad y(t, 0) = x(t), \qquad \forall\, t \in [t_1, t_2], \tag{2.24.2}$$

$$(ii) \quad y(t_1, \varepsilon) = \xi_1, \quad y(t_2, \varepsilon) = \xi_2, \qquad \forall\, \varepsilon \in (-1, 1), \tag{2.24.3}$$

The function y is said a "variation of $\mathbf{x}$" inside $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ parameterized by $\varepsilon \in (-1, 1)$. The set of all variations will be denoted by $\mathcal{V}_{\mathbf{x}}$.

More generally, if $\mathcal{M}$ is a subset of $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ we shall denote $\mathcal{V}_{\mathbf{x}}(\mathcal{M})$ the set of the variations of $\mathbf{x}$ such that, $\forall\, \varepsilon \in (-1, 1)$, the function

$$t \to y_\varepsilon(t) = y(t, \varepsilon), \qquad t \in [t_1, t_2] \tag{2.24.4}$$

is in $\mathcal{M}$.

Observations.
(1) We can imagine that a varied motion y is a bundle of motions with equal initial and final data (see Fig. 2.13).

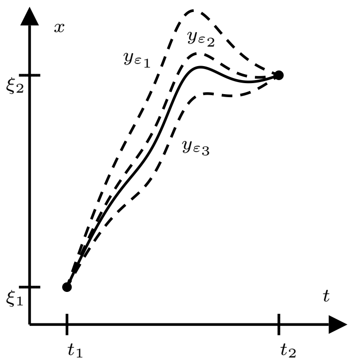


Fig.2.13: Illustration of the variations (dashed curves) of a motion $\mathbf{x}$ (solid curve).

(2) Occasionally it will be useful to think of a variation of $\mathbf{x} \in \mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ as a "regular curve" in the space $\mathcal{M}$: for every $\varepsilon \in (-1, 1)$ one has a point $\mathbf{y}_\varepsilon \in \mathcal{M}$ and $y_0 = \mathbf{x}$ [see Eq. (2.24.4)].

---
[24] For the origin of this name, see the remarks on p. 164 and 241

(3) If F is a function on $\mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ and $y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M})$, it will make sense to consider the function of $\varepsilon \in (-1, 1) : \varepsilon \to F(\mathbf{y}_\varepsilon)$, "value of F along the curve y through $\mathbf{x}$ in the point parameterized by ε".

It is now possible to give a precise definition of stationarity.

23 Definition. *Let $\mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ and let A be a function on $\mathcal{M}$ having the form of Eq. (2.24.1). We shall say that $\mathbf{x} \in \mathcal{M}$ is a "stationarity point" for A in $\mathcal{M}$, if for every $y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M})$ the function [see Eq. (2.24.4)]*

$$\varepsilon \to A(\mathbf{y}_\varepsilon), \qquad \varepsilon \in (-1, 1) \tag{2.24.5}$$

has a stationarity point in $\varepsilon = 0$, i.e.,

$$\frac{d}{d\varepsilon} A(\mathbf{y}_\varepsilon)\big|_{\varepsilon=0} = 0. \qquad \forall\, y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M}). \tag{2.24.6}$$

Observations.

(1) In other words, $\mathbf{x}$ is a stationarity point for A in $\mathcal{M}$ if on every regular curve y through $\mathbf{x}$, the function A, thought of as a function of the parameter ε parameterizing the curve, has a stationarity point in $\varepsilon = 0$, i.e., "in $\mathbf{x}$".
(2) In the theory of maxima and minima of functions $F \in C^\infty(\mathcal{R}^d)$, there are various equivalent definitions of the stationarity points; for instance,
(a) $\frac{\partial F}{\partial x_i}(Vx) = 0$, $i = 1, 2, \ldots, d$.
(b) On every C^∞ curve $\varepsilon \to \mathbf{y}_\varepsilon$, $\varepsilon \in (-1, 1)$, through $\mathbf{x} \equiv \mathbf{y}_0$, the function $F \to F(\mathbf{y}_\varepsilon)$ has zero derivative with respect to ε in $\varepsilon = 0$.
Definition (b) is the "finite-dimensional" analogue inspiring Definition 23: intuitively, one can think of $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ as a vector with infinitely many components $x_t \equiv x(t)$, $t \in [t_1, t_2]$, not independent, however, since they are constrained by the condition that $t \to x_t$ is in $C^\infty([t_1, t_2])$ and $x_{t_1} = \xi_1, x_{t_2} = \xi_2$.
(3) Strictly speaking, one should prove that Eq. (2.24.6) makes sense, i.e., that $\varepsilon \to A(\mathbf{y}_\varepsilon)$ is differentiable in ε. But this is an immediate consequence of the differentiation rules for integrals. Actually, it is easy to find explicit expressions for the derivatives of A. For instance, from Eqs. (2.24.1) and (2.24.5), it follows that

$$\begin{aligned} \frac{d}{d\varepsilon} A(\mathbf{y}_\varepsilon) &= \frac{d}{d\varepsilon} \int_{t_1}^{T_2} \mathcal{L}\big(\frac{\partial y}{\partial t}(t, \varepsilon), y(t, \varepsilon), t\big)\, dt \\ &= \int_{t_1}^{T_2} dt \Big\{ \frac{\partial \mathcal{L}}{\partial \eta}\big(\frac{\partial y}{\partial t}(t, \varepsilon), y(t, \varepsilon), t\big) \cdot \frac{\partial^2 y}{\partial \varepsilon \partial t}(t, \varepsilon)\,\big) \\ &+ \frac{\partial \mathcal{L}}{\partial \xi}\big(\frac{\partial y}{\partial t}(t, \varepsilon), y(t, \varepsilon), t\big) \cdot \frac{\partial y}{\partial \varepsilon}(t, \varepsilon)\Big\}, \end{aligned} \tag{2.24.7}$$

and shortening the notations for $\frac{\partial y}{\partial t}(t, \varepsilon)$ in $\frac{\partial y}{\partial t}$ and for $\frac{\partial^2 y}{\partial \varepsilon \partial t}(t, \varepsilon)$ in $\frac{\partial^2 y}{\partial \varepsilon \partial t}$, and for $y(t, \varepsilon)$ in y, etc., Eq. (2.24.7) can be rewritten:

$$\frac{d}{d\varepsilon}A(\mathbf{y}_\varepsilon) = \int_{t_1}^{t_2} dt \Big\{ \frac{\partial \mathcal{L}}{\partial \eta}\big(\frac{\partial y}{\partial t}, y, t\big)\frac{\partial^2 y}{\partial \varepsilon \partial t} + \frac{\partial \mathcal{L}}{\partial \xi}\big(\frac{\partial y}{\partial t}, y, t\big)\frac{\partial y}{\partial \varepsilon}\Big\}. \tag{2.24.8}$$

Avoiding explicit indication of the arguments $\partial y/\partial t, y, t$ in $\mathcal{L}$ and in its derivatives, a straightforward computation yields

$$\begin{aligned} \frac{d^2}{d\varepsilon^2}A(\mathbf{y}_\varepsilon) &= \int_{t_1}^{t_2} dt \Big\{ \frac{\partial^2 \mathcal{L}}{\partial \eta^2}\Big(\frac{\partial^2 y}{\partial \varepsilon \partial t}\Big)^2 + \frac{\partial^2 \mathcal{L}}{\partial \xi \partial \eta}\frac{\partial^2 y}{\partial \varepsilon \partial t}\frac{\partial y}{\partial \varepsilon} \\ \frac{\partial \mathcal{L}}{\partial \eta}\frac{\partial^3 y}{\partial \varepsilon^2 \partial t} &+ \frac{\partial^2 \mathcal{L}}{\partial \eta \partial \varepsilon}\frac{\partial^2 y}{\partial \varepsilon \partial t}\frac{\partial y}{\partial \varepsilon} + \frac{\partial^2 \mathcal{L}}{\partial \xi^2}\big(\frac{\partial y}{\partial \varepsilon}\big)^2 + \frac{\partial \mathcal{L}}{\partial \xi}\frac{\partial^2 y}{\partial \varepsilon^2}\Big\} \end{aligned} \tag{2.24.9}$$

The higher derivatives could be evaluated with similar procedures; i.e., $\varepsilon \to A(\mathbf{y}_\varepsilon)$ is a C^∞ function.

As in the case of the functions on $\mathcal{R}^d$, it is convenient to distinguish between stationary points and points of "local" or "relative" minimum.

24 Definition. *If $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$, we say that $\mathbf{x}$ is a"local" minimum for A defined by Eq. (2.24.1) on $\mathcal{M}$ if for all varied motions $y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M})$, the function $\varepsilon \to A(\mathbf{y}_\varepsilon)$ [see Eq. (2.24.5)] has a relative minimum in $\varepsilon = 0$.*

Observations.
(1) A has a local minimum in $\mathbf{x}$ on $\mathcal{M}$ if on every regular curve y through $\mathbf{x}$ lying on $\mathcal{M}$, it has a local minimum in $\mathbf{x}$.
(2) A necessary condition for A to have a local minimum relative to $\mathcal{M}$ in $\mathbf{x} \in \mathcal{M}$ is that $\mathbf{x}$ is a stationarity point for A on $\mathcal{M}$.
(3) A necessary condition in order that a stationarity point for A on $\mathcal{M}$ is a local minimum on $\mathcal{M}$ is that

$$\frac{d^2}{d\varepsilon^2}A(\mathbf{y}_\varepsilon)\Big|_{\varepsilon=0} \geq 0, \qquad \forall\, y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M}) \tag{2.24.10}$$

if $\mathbf{x}$ is the point of stationarity.
(4) If $\mathbf{x} \in \mathcal{M}$ is an absolute minimum point for A on $\mathcal{M}$, i.e., if $A(\mathbf{x}') > A(\mathbf{x}), \forall\, \mathbf{x}' \in \mathcal{M}$, then $\mathbf{x}$ is also a local minimum point for A on $\mathcal{M}$.
(5) If A has a local minimum in $\mathbf{x}$ relative to $\mathcal{M}$ it must be that, given $y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M})$, there is $\eta > 0$ such that if $\varepsilon \in [-\eta, \eta]$, then $A(\mathbf{y}_\varepsilon) \geq A(\mathbf{x})$: this value of η may, however, depend on the choice of y.
(6) One could be tempted to define a local minimum by requiring that $A(\mathbf{x}) \leq A(\mathbf{x}'), \forall\, \mathbf{x}' \in \mathcal{M}$ and "close enough" to $\mathbf{x}$. But the meaning of "close enough" would be unclear.

A necessary and sufficient stationarity criterion, which is as "simple" as the one usually considered in the case of the stationarity of functions on $\mathcal{R}^d$ and concerning the vanishing of the gradient (see Observation 2 (a), to Definition 23), is the following.

36 Proposition. *The motion* $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ *is a stationary point for Eq. (2.24.1) on all of* $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ *if and only if*

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \eta}(\dot{x}(t), x(t), t) = \frac{\partial \mathcal{L}}{\partial \xi}(\dot{x}(t), x(t), t), \qquad \forall\, t \in [t_1, t_2]. \tag{2.24.11}$$

Observations.
(1) In this proposition, it is essential that the set $\mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ on which stationarity is considered coincides with $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ itself.
(2) Equation (2.24.11) can be thought of as a differential equation for the function $t \to x(t)$, $t \in [t,t_2]$, i.e., as an equation for the determination of the stationarity points of A on the entire set $\mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$. When Eq. (2.24.11) is viewed in this way, it is called the "Euler-Lagrange" equation for A or $\mathcal{L}$. As it emerges from the proof, it is analogous to the condition of vanishing in the stationarity problem for functions on $\mathcal{R}^d$ (see observation 2 (a) to Definition 23, p.128).
(3) It has to be kept in mind that, in general, Eq. (2.24.11) is not a differential equation in the sense of Definition 1, p.14: in important cases, however, Eq. (2.24.11) is equivalent to a differential equation in that sense (see Problems 4-6 at the end of this section).

Proof. It reduces to a check. Let $y \in \mathcal{V}_{\mathbf{x}}$ and set

$$z(t) = \frac{\partial y}{\partial t}(t, 0), \qquad t \in [t_1, t_2], \tag{2.24.12}$$

$$\dot{z}(t) = \frac{\partial^2 y}{\partial t \partial \varepsilon}(t, 0), \qquad t \in [t_1, t_2] \tag{2.24.13}$$

and note that Eq. (2.24.2) implies, $\forall\, t \in [t_1, t_2]$:

$$y(t, 0) = x(t), \qquad \frac{\partial y}{\partial t}(t, 0) = \dot{x}(t) \tag{2.24.14}$$

while Eq. (2.24.3) gives

$$z(t_1) = z(t_2) = 0 \tag{2.24.15}$$

Then, with the above notations, Eq. (2.24.8) becomes

$$\frac{d^2}{d\varepsilon^2} A(\mathbf{y}_\varepsilon)\Big|_{\varepsilon=0} = \int_{t_1}^{t_2} dt \Big\{ \frac{\partial \mathcal{L}}{\partial \eta}(\dot{x}(t), x(t), t)\dot{z}(t) + \frac{\partial \mathcal{L}}{\partial \xi}(x(t), x(t), t) z(t)) \Big\}. \tag{2.24.16}$$

As $\mathbf{y}_\varepsilon$ varies in $\mathcal{V}_{\mathbf{x}}$, the function z defined by Eq. (2.24.12) spans the entire set $\mathcal{M}_{t_1,t_2}(0, 0)$. In fact, Eq. (2.24.15) shows that $z \in \mathcal{M}_{t_1,t_2}(0, 0)$; furthermore, given arbitrarily $\overline{z} \in \mathcal{M}_{t_1,t_2}(0, 0)$ and setting

$$\overline{y}(t,\varepsilon) = x(t) + \varepsilon\overline{z}(t) \tag{2.24.17}$$

for $\varepsilon \in (-1,1)$, $t \in [t_1,t_2]$, one constructs a varied motion $y \in \mathcal{V}_{\mathbf{x}}$, which, via Eq. (2.24.12), exactly generates $\overline{z}$.

The wide arbitrariness of $\mathbf{z}$ in Eq. (2.24.16) can then be used to deduce conditions on $\mathbf{x}$. For this purpose it is convenient to eliminate $\dot{z}(t)$ from Eq. (2.24.16) by integrating the first term by parts; one finds:

$$\begin{aligned} \frac{d^2}{d\varepsilon^2} A(\mathbf{y}_\varepsilon)\Big|_{\varepsilon=0} &= \Big[\frac{\partial \mathcal{L}}{\partial \eta}(\dot{x}(t), x(t), t) z(t)\Big]_{t_1}^{t_2} \\ &- \int_{t_1}^{t_2} \Big\{ \frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \eta}(\dot{x}(t), x(t), t)\Big) - \frac{\partial \mathcal{L}}{\partial \xi}(\dot{x}(t), x(t), t)\Big\} z(t)\, dt \end{aligned} \tag{2.24.18}$$

which, by Eq. (2.24.15) and by the preceding remark on the arbitrariness of z, shows that $(dA/d\varepsilon)(\mathbf{y}_\varepsilon)|_{\varepsilon=0} = 0$, $\forall\, y \in \mathcal{V}_{\mathbf{x}}$, becomes:

$$0 = \int_{t_1}^{t_2} \Big\{ \frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \eta}(\dot{x}(t), x(t), t)\Big) - \frac{\partial \mathcal{L}}{\partial \xi}(\dot{x}(t), x(t), t)\Big\} z(t)\, dt, \tag{2.24.19}$$

$\forall\, z \in \mathcal{M}_{t_1,t_2}(0,0)$. The equivalence between Eqs. (2.24.19) and (2.24.11) is implied by the principle of vanishing integrals (see Appendix D). mbe

As a consequence of Proposition 36, it is possible to answer the question raised at the beginning of this section. In fact, if one defines for $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1,\xi_2)$.

$$A(\mathbf{x}) = \int_{t_1}^{t_2} \big(\frac{1}{2} m\dot{x}(t)^2 - V(x(t))\big)\, dt \tag{2.24.20}$$

the following proposition holds.

37 Proposition. *The motion* $\mathbf{x}$ *of a point, with mass* $m > 0$ *developing from* ξ_1 *to* ξ_2 *in the time interval* $[t_1,t_2]$ *under the influence of a force with potential energy* $V \in C^\infty(\mathcal{R})$*, makes the action of Eq. (2.24.20) on* $\mathcal{M}_{t_1,t_2}(\xi_1,\xi_2)$ *stationary, i. e., it makes stationary the action with Lagrangian density*

$$\mathcal{L}(\eta,\xi,t) = \frac{1}{2} m\eta^2 - V(\xi). \tag{2.24.21}$$

Proof. In fact, Eq. (2.24.11) becomes

$$\frac{d}{dt} m\,\dot{x}(t) = -\frac{\partial V}{\partial \xi}(x(t)), \qquad t \in [t_1,t_2] \tag{2.24.22}$$

which is the equation of motion. mbe

Furthermore, the following interesting proposition holds.

38 Proposition. *Let $t \to \overline{x}(t)$, $t \in \mathcal{R}$, be a motion of a point mass with $m > 0$ developing under the action of a force with potential energy $V \in C^\infty(\mathcal{R})$, bounded from below. Given $t_1 \in \mathcal{R}$, there exists $\overline{t} > t_1$ such that if $t_2 \in [t_1, \overline{t}]$ the motion $t \to \overline{x}(t)$ observed for $t \in [t_1, t_2]$, i.e., as an element of $\mathcal{M}_{t_1,t_2}(\overline{x}(t_1), \overline{x}(t_2))$, not only is a stationarity point for the action with Lagrangian Eq. (2.24.21), but is also a local minimum for it in $\mathcal{M}_{t_1,t_2}(\overline{x}(t_1), \overline{x}(t_2))$.*

Observation. The proposition motivates the name "principle of the least action" occasionally given to the Propositions 37 and 38.

Proof. By the observation 5 to Definition 24, p.129, given $y \in \mathcal{V}_{\overline{\mathbf{x}}}$, we must find a η_y such that $A(\mathbf{y}_\varepsilon) \geq A(\overline{\mathbf{x}})$, $\forall \varepsilon \in [-\eta_y, \eta_y]$.

Given $t_2 > t_1$ and $y \in \mathcal{M}_{t_1,t_2}(\overline{x}(t_1), \overline{x}(t_2))$, define η_y so that $|y_\varepsilon(t) - \overline{x}(t)| \leq 1, \forall t \in [t_1, t_2]$, $\forall \varepsilon \in [-\eta_y, \eta_y]$. The comparison of $A(\mathbf{y}_\varepsilon)$ with $A(\overline{\mathbf{x}})$ yields

$$\begin{aligned} A(\mathbf{y}_\varepsilon) - A(\overline{\mathbf{x}}) = \int_{t_1}^{t_2} \Big\{ & \frac{m}{2}\big((\dot{x}(t) + \dot{z}(t))^2 \dot{\overline{x}}(t)^2\big) \\ & - \big(V(\overline{x}(t) + z(t)) - V(\overline{x}(t))\big)\Big\}\, dt, \end{aligned} \tag{2.24.23}$$

where we set $z(t) = y_\varepsilon(t) - \overline{x}(t)$, $t \in [t_1, t_2]$. This is a function z which has the property

$$z(t_1) = z(t_2) = 0 \tag{2.24.24}$$

and it is ε dependent. To show that Eq. (2.24.23) is > 0, apply the Taylor-Lagrange formula (see Appendix B):

$$V(\xi') - V(\xi) = \frac{\partial V}{\partial \xi}(\xi)\,(\xi' - \xi) + \varphi(\xi', \xi)\frac{(\xi' - \xi)^2}{2}, \tag{2.24.25}$$

where $\varphi \in C^\infty(\mathcal{R}^2)$ is a suitable function. Then Eq. (2.24.23) becomes

$$\begin{aligned} A(\mathbf{y}_\varepsilon) - A(\overline{\mathbf{x}}) = \int_{t_1}^{t_2} \Big\{ & \Big[m\frac{\dot{z}(t)^2}{2} - \varphi(\overline{x}(t) + z(t))\frac{z(t)^2}{2}\Big] \\ & + \Big[m\dot{\overline{x}}(t)\, z(t) - \frac{\partial V}{\partial \xi}(\overline{x}(t))z(t)\Big]\Big\}\, dt \end{aligned} \tag{2.24.26}$$

Integrating the first term in the second set of square brackets by parts and using the equation of motion for $\overline{x}$, Eqs. (2.24.22) and (2.24.24), one realizes that the integral of the term within the second set of square brackets in Eq. (2.24.26) vanishes. Therefore, if

$$M = \max_{\substack{t \in [t_1, t_1+1] \\ |\zeta| \leq 1}} |\varphi(\overline{x}(t) + \zeta, \overline{x}(t))|, \tag{2.24.27}$$

one sees that, if $|\varepsilon| < \eta_y$, Eq. (2.24.26) implies

$$A(\mathbf{y}_\varepsilon) - A(\overline{\mathbf{x}}) \geq \frac{m}{2}\int_{t_1}^{t_2} \dot{z}(t)^2\,dt - \frac{M}{2}\int_{t_1}^{t_2} z(t)^2\,dt; \tag{2.24.28}$$

if $t_2 - t_1 < 1$, which is a condition that can be implemented by supposing $t_2 < \overline{t}$ and, without loss of generality,

$$\overline{t} < t_1 + 1. \tag{2.24.29}$$

On the other hand, since $z(t_1) = 0$,

$$z(t) = \int_{t_1}^{t} \dot{z}(\tau)\,d\tau, \tag{2.24.30}$$

and applying the Cauchy-Schwartz inequality (see Appendix A), which generally looks like, $\forall\, f, g \in C^\infty([t_1, t_2])$,

$$\Big|\int_{t_1}^{t_2} f(\tau)g(\tau)d\tau\Big| \leq \Big(\int_{t_1}^{t_2} f(\tau)^2 d\tau\Big)^{\frac{1}{2}}\Big(\int_{t_1}^{t_2} g(\tau)^2 d\tau\Big)^{\frac{1}{2}} \tag{2.24.31}$$

one finds,

$$\begin{aligned}\int_{t_1}^{t_2} z(t)^2\,dt &= \int_{t_1}^{t_2} \Big|\int_{t_1}^{t} \dot{z}(\tau)\cdot 1\,d\tau\Big|^2 dt \leq \int_{t_1}^{t_2} dt\Big(\int_{t_1}^{t} \dot{z}(\tau)^2 d\tau\Big)\Big(\int_{t_1}^{t} 1 d\tau\Big)\\ &\leq \int_{t_1}^{t_2} dt(t-t_1)\Big(\int_{t_1}^{t_2} \dot{z}(\tau)^2 d\tau\Big) = \frac{(t_2-t_1)^2}{2}\int_{t_1}^{t_2} \dot{z}(\tau)^2 d\tau\end{aligned} \tag{2.24.32}$$

from Eq. (2.24.30). Hence Eqs. (2.24.28) and (2.24.32) mean

$$A(\mathbf{y}_\varepsilon) - A(\overline{\mathbf{x}}) \geq \frac{1}{2}\Big(m - \frac{M}{2}(t_2-t_1)^2\Big)\int_{t_1}^{t_2} \dot{z}(\tau)^2 d\tau \tag{2.24.33}$$

which implies $A(\mathbf{y}_\varepsilon) - A(\overline{\mathbf{x}}) > 0$ if $t_2 \in [t_1, \overline{t}]$ and if $\overline{t}$ is close enough to t_1 (precisely so that $\overline{t}-t_1 < 1$ and $2m - M(\overline{t}-t_1)^2 > 0$), $\forall\, y \in \mathcal{V}_{\mathbf{x}}, \forall\, \varepsilon \in [-\eta_y, \eta_y]$.
mbe

In the context of Proposition 38, one can wonder about what happens when the interval $[t_1, t_2]$ is not small: and one realizes that it is always possible to cut the interval $[t_1, t_2]$ into finitely many small intervals such that the action is locally minimal on the variations of the restrictions of $\overline{\mathbf{x}}$ to such intervals.

This situation is strongly reminiscent of the properties of the geodesics on curved surfaces. For instance, on a sphere, a line joining two points along a great circle ("geodesic of the sphere") has the property of being the line shortest among all those joining the two points and lying on the sphere, provided their distance, measured along the line itself, is small enough. However, if the two points are not close enough, it is generally no longer true that such a line is the shortest ("close enough" here means closer than πR if R is the radius).

Finally, let us meditate upon the following important comment: we wish to stress the fact that the stationarity (or minimality) of A is an "intrinsic property", i.e., it is independent of the way the motion is described. To make this precise, let $\xi \to \gamma(\xi)$ be a C^∞ function defining a "nonsingular" change of variables (i.e., such that $\gamma'(\xi) \equiv \frac{d\gamma}{d\xi}(\xi) \neq 0$). We can then use as a coordinate for the point ξ the quantity $\gamma(\xi)$.

Let Γ be the inverse function to γ defined on the open interval $I = \gamma(\mathcal{R}) = \gamma$-image of $\mathcal{R}$. Suppose, for simplicity, $\gamma(\mathcal{R}) = I = \mathcal{R}$.

A motion in $\mathcal{R}$, $t \to x(t)$, $t \in [t_1, t_2]$, can be described by the function $t \to s(t) = \gamma(x(t))$, $t \in [t_1, t_2]$. We shall say that such a function describes the motion x in the system of coordinates on $\mathcal{R}$ associated with the function γ. There is a one-to-one correspondence B between motions $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ and motions $\mathbf{s} \in \mathcal{M}_{t_1,t_2}(\gamma(\xi_1), \gamma(\xi_2))$: it is established by the relations

$$s(t) = \gamma(x(t)), \quad x(t) = \Gamma(s(t)), \qquad t \in [t_1, t_2] \tag{2.24.34}$$

The correspondence of Eq. (2.24.34) will be denoted by $\mathbf{s} \stackrel{def}{=} B\mathbf{x}$. Let Γ' be the derivative of Γ, then $\mathbf{s} = B\mathbf{x}$ implies

$$\dot{x}(t) = \Gamma'(s(t))\, \dot{s}(t), \tag{2.24.35}$$

and it has to be remarked that the Lagrangians

$$\mathcal{L}(\eta, \xi, t) = \frac{m}{2}\eta^2 - V(\xi), \tag{2.24.36}$$

$$\widetilde{\mathcal{L}}(\eta, \xi, t) = \frac{m\Gamma'(\xi)^2}{2}\eta^2 - V(\Gamma(\xi)), \tag{2.24.37}$$

attribute the same action to the motions $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1, \xi_2)$ and, respectively, $\mathbf{s} \in \mathcal{M}_{t_1,t_2}(\gamma(\xi_1), \gamma(\xi_2))$; i.e., if $\mathbf{s} = B\mathbf{x}$,

$$\begin{aligned} A(\mathbf{x}) &= \int_{t_1}^{t_2} dt\, \big(\frac{m\dot{x}(t)^2}{2} - V(x(t))\big) \\ &\equiv \widetilde{A}(\mathbf{s}) = \int_{t_1}^{t_2} dt\, \big(\frac{m\Gamma'(s(t))^2 \dot{s}(t)^2}{2} - V(\Gamma(s(t)))\big)\, dt \end{aligned} \tag{2.24.38}$$

which follows from Eqs. (2.24.34) and (2.24.35).

If $y \in \mathcal{V}_{\mathbf{x}}(\mathcal{M})$ it is natural to associate with y the element $B\,y \in \mathcal{V}_{\mathbf{s}}(B\mathcal{M})$

$$(By)(t, \varepsilon) \stackrel{def}{=} \gamma(y(t, \varepsilon)), \qquad (t, \varepsilon) \in [t_1, t_2] \times (-1, 1) \tag{2.24.39}$$

and $B\mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\gamma(\xi_1), \gamma(\xi_2))$ is the image of $\mathcal{M}$ via the map of Eq. (2.24.39).

It is then an immediate consequence of Definitions 23 and 24 that if A is stationary or locally minimal on $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\xi_1), \xi_2))$ in $\mathcal{M}$, then $\widetilde{A}$ also is

stationary or locally minimal on $\mathbf{s} = B\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\gamma(\xi_1), \gamma(\xi_2))$ in $B\mathcal{M}$ and vice versa. In particular, this means that the equations

$$\begin{aligned} \frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \eta}(\dot{x}(t), x(t), t)\Big) &= \frac{\partial \mathcal{L}}{\partial \xi}(\dot{x}(t), x(t), t) \\ \frac{d}{dt}\Big(\frac{\partial \widetilde{\mathcal{L}}}{\partial \eta}(\dot{s}(t), s(t), t)\Big) &= \frac{\partial \widetilde{\mathcal{L}}}{\partial \xi}(\dot{s}(t), s(t), t) \end{aligned} \tag{2.24.40}$$

are "equivalent" if $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ are given by Eqs. (2.24.36) and (2.24.37).

As we shall see, this invariance property of the stationarity (or of the local minimality) with respect to changes of coordinates is perhaps the most interesting aspect of all the considerations of this section. We shall meet some of its very remarkable applications in the theory of systems with many degrees of freedom.

Concluding Remarks

(1) In the analysis of this section we always dealt with conservative systems. In fact, it is not possible to give a simple formulation of the stationary action principle for dissipative motions without introducing singular Lagrangians (see Problems 12-15 at the end of this section).
(2) The action of a motion x with Lagrangian (2.24.36) can be thought of as the product of $(t_2 - t_1)$ times the difference between the average value, in $[t_1, t_2]$, of the kinetic energy and the average value of the potential energy:

$$\frac{A(\mathbf{x})}{t_2 - t_1} = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \frac{m\,\dot{x}(t)^2}{2}\,dt - \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} V(x(t))\,dt. \tag{2.24.41}$$

It is for this reason that one can say that the motion developing for $t \in [t_1, t_2]$ between t_1, and t_2 under the influence of a force of given potential energy V is the one that minimizes the difference between the average kinetic energy and the average potential energy in every short enough time interval in $[t_1, t_2]$.

We leave it to the reader to elaborate his own philosophical considerations on this beautiful mathematical property. The interested reader could go through the history of the variational principles in mechanics and, more generally, in physics, to understand how subjective considerations (as we would call them today) have influenced the formulation of the variational principles themselves and the recognition of their equivalence to the Newtonian equations of motion; see also the comments on p. 164 and p. 242 and the Euler's quotation at the beginning of this section.

2.24.1 Exercises and Problems

1. Compute the action between $t_1 = 0$ and $t_2 = 2\pi/\omega$ of the motions of an harmonic oscillator with mass $m > 0$ and pulsation ω.

2. Same as Problem 1 with t_2 arbitrary ($t_2 \neq 2\pi/\omega$).

3. Compute the action between $t_1 = 0$ and arbitrary t_2 of the motions of a point mass with mass $m > 0$ subject to the force $f = -mg, g > 0$.

4. Let $\mathcal{L} \in C^\infty(\mathcal{R}^2)$ be such that the correspondence $(\eta, \xi) \to (\frac{\partial \mathcal{L}}{\partial \eta}(p, \xi), \xi)$ can be inverted in class C^∞ as a mapping of $\mathcal{R}^2$ onto $\mathcal{R}^2$ and let $(p, \xi) \to (f(p, \xi), \xi)$ be the inverse map. Set $H(p, \xi) = pf(p, \xi) - \mathcal{L}(f(p, \xi), \xi) \equiv [p\eta - \mathcal{L}(\eta, \xi)]_{\eta = f(p,\xi)}$, and check that the "Lagrange equations"

$$\dot{\xi} = \eta, \qquad \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \eta}(\eta, \xi) = \frac{\partial \mathcal{L}}{\partial \xi}(\eta, \xi)$$

are equivalent to the "Hamilton equations"

$$\dot{p} = -\frac{\partial H}{\partial \xi}(p, \xi), \qquad \dot{\xi} = \frac{\partial H}{\partial p}(p, \xi).$$

The motion described in terms of p and ξ, $t \to (p(t), \xi(t))$, is a solution of this differential equation and any of its solutions is called a "Hamiltonian motion" and the space $\mathcal{R}^2$, thought of as the space of the initial data for the above equations, is called a "phase space". (*Hint:* Note that since, by definition of p and f, one has $p \equiv \frac{\partial \mathcal{L}}{\partial \eta}(f(p, \xi), \xi)$, it follows that

$$\frac{\partial H}{\partial p}(p, \xi) = f(p, \xi) + p\frac{\partial f}{\partial p}(p, \xi) - \frac{\partial \mathcal{L}}{\partial \eta}(f(p, \xi), \xi)\frac{\partial f}{\partial p}(p, \xi) \equiv f(p, \xi) = \eta$$

and

$$\begin{aligned}\frac{\partial H}{\partial \xi}(p, \xi) =& p\frac{\partial f}{d\xi}(p, \xi) - \frac{\partial \mathcal{L}}{\partial \eta}(f(p, \xi), \xi)\frac{\partial f}{d\xi}(p, \xi) - \frac{\partial \mathcal{L}}{\partial \xi}(f(p, \xi), \xi) \\ \equiv& -\frac{\partial \mathcal{L}}{\partial \xi}(f(p, \xi), \xi) \equiv -\frac{\partial \mathcal{L}}{\partial \xi}(\eta, \xi)\end{aligned}$$

having used the definition of H.)

5. The function H in Problem 4 can be expressed in terms of $\mathcal{L}$ and vice versa as

$$H(p, \xi) = \max_{\eta \in \mathcal{R}}(p\eta - \mathcal{L}(\eta, \xi)), \qquad \mathcal{L}(\eta, \xi) = \max_{p \in \mathcal{R}}(p\eta - H(p, \xi))$$

("Legendre duality"), if the maximum is attained at a unique point $\overline{\eta}$ or $\overline{p}$, respectively, and, furthermore, if $\overline{\eta}, \overline{p}$ are the only stationarity points of the functions in brackets as functions of η or p, respectively. (*Hint:* Write the stationarity conditions for $p\eta - \mathcal{L}(\eta, \xi)$ and those for $p\eta - H(p, \xi)$ with respect to η or, respectively, to p. Then use the definition of H in Problem 4.)

6. The "Hamilton equations" $\dot{p} = -\frac{\partial H}{\partial \xi}(p, \xi)$, $\dot{\xi} = \frac{\partial H}{\partial p}(p, \xi)$. with Hamiltonian $H \in C^\infty(\mathcal{R}^2)$ can be obtained by imposing stationarity of

$$S = \int_{t_1}^{t_2} (p(t)\dot{x}(t) - H(p(t), x(t)))\, dt$$

in the space $\mathcal{M}_{t_1,t_2}\Big((\pi_1, \xi_1), (\pi_2, \xi_2)\Big)$ of the $C^\infty([t_1, t_2])$ functions $t \to (p(t), q(t)) \in \mathcal{R}^2$ such that $p(t_1) = \pi_1, p(t_2) = \pi_2, x(t_1) = \xi_1, x(t_2) = \xi_2$, ("Hamilton's principle"). (*Hint:* Apply Proposition 36, Eq. (2.24.1), with $\mathcal{L}(\dot{p}, \dot{x}, p, x) = p\dot{x} - H(p, x)$.)

7. In the context of Problem 6, show that the same Hamilton equations can be obtained by imposing stationarity of S on the larger space $\widetilde{\mathcal{M}}_{t_1,t_2}(\xi_1, \xi_2)$ of the $C^\infty([t_1, t_2])$ functions $t \to (p(t), q(t)) \in \mathcal{R}^2$ such that $x(t_1) = \xi_1, x(t_2) = \xi_2$. (*Hint:* Go through the proof of Proposition 36 using the special form of the Lagrangian $\mathcal{L}(\dot{p}, \dot{x}, p, x) = p\dot{x} - H(p, x)$.)

8. Let $t \to (p(t), x(t)) \in \mathcal{R}^2$ be a motion verifying the Hamilton equations of Problems 4 and 6. Show that the quantity S defined in Problem 6 coincides with $\int_{t_1}^{t_2} \mathcal{L}(\dot{x}(t), x(t))\, dt$, i.e., with the action of the same motion (of course, if $\mathcal{L}$ and H are related as in Problem 4).

9. Extend Problems 4 and 7 to the case when H and $\mathcal{L}$ depend explicitly on time.

10.* Let H be as in Problem 4 and let $S_t(p, x) = (p(t), x(t)), t \ge 0$, be the solution of the Hamilton equations (as in Problem 14), supposed normal, with (p, x) as initial datum at $t = 0$. Let $A \subset \mathcal{R}^2$ be a (Riemann) measurable region. Show that $\text{area}(S_t A) = \text{area}(A)$, $\forall\, t \ge 0$, if $S_t A = \{\text{set of points of the form } S_t(p, x), \text{ with } (p, x) \in A\}$ ("Liouville's theorem"). (*Hint:* In general, let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ be an autonomous normal differential equation in $\mathcal{R}^d$. Set $\mathbf{y} = S_t \mathbf{x}$, for $t \ge 0$. Then

$$\text{volume}(S_t A) = \int_{S_t A} d\mathbf{x} = \int_A \left|\det\left(\frac{\partial S_{-t}(\mathbf{y})}{d\mathbf{y}}\right)\right| d\mathbf{y}$$

where $\partial S_t(\mathbf{y})/\partial \mathbf{y}$ denotes the Jacobian matrix of the coordinate transformation $\mathbf{x} = S_t(\mathbf{y})$. This formula shows that if $\det\left(\frac{\partial S_{-t}(\mathbf{y})}{\partial \mathbf{y}}\right) > 0$, the modulus symbol is irrelevant and $t \to \text{volume}(S_t A)$ is a C^∞ function, and

$$\frac{d}{dt}\text{volume}(S_t A)|_{t=\tau} = \int_A \left[\frac{d}{dt}\det\left(\frac{\partial S_{-t}(\mathbf{y})}{\partial \mathbf{y}}\right)\right]_{t=\tau} d\mathbf{y}.$$

But (see §2.6) $S_{t+\tau} = S_t S_\tau$, hence, the last expression is equal to:

$$\int_A \left[\frac{d}{dt}\det\left(\frac{\partial S_{-t-\tau}(\mathbf{y})}{\partial \mathbf{y}}\right)\right]_{t=0} d\mathbf{y} = \int_A \left[\frac{d}{dt}\det\left(\frac{\partial S_{-t}(S_{-\tau}(\mathbf{y}))}{\partial \mathbf{y}}\right)\right]_{t=0} d\mathbf{y}$$
$$= \int_A \left[\frac{d}{dt}\det\left(\frac{\partial S_{-t}(S_{-\tau}(\mathbf{y}))}{\partial S_{-\tau}(\mathbf{y})}\right)\right]_{t=0} \det\left(\frac{\partial S_{-\tau}(\mathbf{y})}{\partial \mathbf{y}}\right) d\mathbf{y}$$

by the composite function differentiation rule and by the determinant rules.

It is then sufficient to check that, under suitable circumstances, the derivative

$$\left[\frac{d}{dt}\det\left(\frac{\partial S_{-t}(\mathbf{x})}{\partial \mathbf{x}}\right)\right]_{t=0} \equiv 0, \qquad \forall\, \mathbf{x} \in \mathcal{R}^d$$

to infer the volume conservation under the same circumstances. If $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, it follows that

$$S_t \mathbf{x} = \mathbf{x} + t\,\mathbf{f}(\mathbf{x}) + t^2\, \boldsymbol{\varphi}(\mathbf{x}, t)$$

by the Taylor-Lagrange formula (see Appendix B), where $\boldsymbol{\varphi}$ is a C^∞ function of $\mathbf{x}$ and t. Hence,

$$\det \frac{\partial S_t(\mathbf{x})}{\partial \mathbf{x}} = \det\left(1 + t\frac{\partial f^{(i)}(\mathbf{x})}{\partial x_j} + t^2\, \frac{\partial \varphi^{(i)}(\mathbf{x}, t)}{\partial x_j}\right);$$

hence, by developing the determinant

$$\det \frac{\partial S_t(\mathbf{x})}{\partial \mathbf{x}} == 1 + t\sum_{j=1}^{d} \frac{\partial f^{(j)}(\mathbf{x})}{\partial x_j} + t^2\, \psi(\mathbf{x}, t),$$

where ψ is a suitable C^∞ function of x, t. Hence, the derivative of $\det(\partial S_t(\mathbf{x})/\partial \mathbf{x})$ for $t = 0$ is $\sum_{j=1}^d \frac{\partial f^{(j)}(\mathbf{x})}{\partial x_j} \equiv \text{div}\, \mathbf{f}(\mathbf{x})$, wherein the right-hand side is the notation used in physics for the left-hand side ("divergence of $\mathbf{f}$").

Therefore, if $\operatorname{div}\mathbf{f} = 0$, the flow S_t generated by $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ preserves the volume (this also motivates the name "divergence" given to $\operatorname{div}\mathbf{f}(\mathbf{x})$ since it measures the rate of increase of volume under the transformation S_t). In fact, it follows from the above considerations that $\det(\partial S_t(\mathbf{x})/\partial\mathbf{x}) \equiv 1$, being constant and equal to 1 for $t = 0$. Then note that the Hamilton equations are divergenceless.)

11.* Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ be an autonomous normal differential equation in $\mathcal{R}^d$, $\mathbf{f} \in C^\infty(\mathcal{R}^d)$, and suppose $\operatorname{div}\mathbf{f}(\mathbf{x}) \equiv \sum_{j=1}^d \frac{\partial f^{(j)}(\mathbf{x})}{\partial x_j} = 0$, $\forall\,\mathbf{x} \in \mathcal{R}^d$. So, by the hint to Problem 10, it follows that the solution flow $(S_t)_{t\in\mathcal{R}_+}$ preserves the volume: $\text{volume}S_tA \equiv \text{volume}A$.
Suppose that the solution flow maps a bounded open set $\Omega \subset \mathcal{R}^d$ into itself: $S_t\Omega \subset \Omega$, $\forall\, t \in \mathcal{R}_+$. Show that given $\mathbf{x}_0 \in \Omega$, $t_0 > 0$, and a neighborhood $U \subset \Omega$ of x_0, there exists $t \ge t_0$ such that $S_tU \cap U \ne \emptyset$; i.e., close to any point $x_0 \in \Omega$, there is another point which comes as close to x_0 after a given, arbitrarily large, time ("Poincaré's recurrence theorem"). (*Hint:* Suppose $S_{t_0}U \cap U \ne \emptyset$, otherwise $t = t_0$; then consider $S_{2t_0}U \cap U \ne \emptyset$, show that the three sets $U, S_{t_0}U, S_{2t_0}U$ must be pairwise disjoint because if $S_{2t_0} \cap U \ne \emptyset$ then $t = 2t_0$. In the first case, consider $S_{3t_0}U$: if $S_{3t_0}U \cap U = \emptyset$ the four sets $U, S_{t_0}, S_{2t_0}U$ and $S_{3t_0}U$ must be pairwise disjoint; if not, take $t = 3t_0$, etc. The result could fail only if the sequence $U, S_{t_0}U, S_{2t_0}U, \ldots, S_{kt_0}U, \ldots$ is an infinite sequence of pairwise disjoint sets. However, in such a case, $\text{volume}(\Omega) \ge \sum_{k=0}^\infty \text{volume}(S_{kt_0}U) = \sum_{k=0}^\infty \text{volume}(U) = +\infty$ because U is open and $\text{volume}(U) > 0$, which is absurd since Ω is a bounded set.)

12. Show that the equation $\ddot{x} + \gamma\dot{x} = 0$, $\gamma > 0$ describing a free particle moving under the action of linear friction is the Euler-Lagrange equation associated with the Lagrangian $\mathcal{L}(\dot{x}, x) = \dot{x}\log\dot{x} - \gamma x$ in the region $\dot{x} > 0$, [27]. (Define the Euler-Lagrange equations by Eq. (2.24.11), i.e. as $(d/dt)(\partial\mathcal{L}/\partial\dot{x}) = \partial\mathcal{L}/\partial x$.)

13. Let $V \in C^\infty(\mathcal{R})$ be bounded below. Show that if $F \in C^\infty(\mathcal{R})$ has a non vanishing derivative, the equations $\ddot{x} = -(dV/dx)(x)$ can be described by the Lagrangian function

$$\mathcal{L}(\eta, \xi) = \eta \int_1^\eta \frac{F(\frac{1}{2}y^2 + V(\xi))}{y^2}\, dy$$

in the region $\dot{x} > 0$, i.e., $\eta > 0$. What does $\mathcal{L}$ become if $F(e) \equiv e$, $\forall\, e \in \mathcal{R}$? Is this consistent with the alternative Lagrangian $\widetilde{\mathcal{L}} = \frac{1}{2}\eta^2 - V(\xi)$? (see [27]).

14. Consider the damped oscillator $\ddot{x} + \dot{x} + \omega^2 x = 0$ and let $\alpha = (4\omega^2 - \gamma)^{-\frac{1}{2}}$, $\gamma > 0$. Show that in the region $\eta > 0, \xi > 0$, the Lagrangian

$$\mathcal{L}(\eta, \xi) = -\frac{1}{2}\log(\eta^2 + \gamma\eta\xi + \omega^2\xi) + \alpha\,(2\frac{\eta}{\xi} + \gamma)\text{arctg}\,(2\frac{\eta}{\xi} + \gamma)$$

has, as Euler-Lagrange equations, the damped oscillator equations (see [27]).

15. Let $\ddot{x} = g(\dot{x}, x)$ be a differential equation. Show that in order that a function $\mathcal{L}$ on a subset A of $\mathcal{R}^2$ generates (via the Euler-Lagrange equations) the equation $\ddot{x} = g(\dot{x}, x)$ for the motions developing in A, it must be

$$\frac{\partial\mathcal{L}}{\partial\eta} - \eta\frac{\partial^2\mathcal{L}}{\partial\xi\partial\eta} - g(\eta, \xi)\frac{\partial^2\mathcal{L}}{\partial\eta^2} = 0, \qquad \forall\,(\eta, \xi) \in A$$

(*Hint*: Write the Euler-Lagrange equations substituting $\ddot{x}$ with $g(\dot{x}, x)$) (see [27]).[25]

[25] The last four problems are taken from [27]. The equation for $\mathcal{L}$ in Problem 15 allows one to find many Lagrangians for the same equation. Note, however, that such Lagrangians will generally be singular somewhere in $\mathcal{R}^2$, always so, probably, if the equation $\ddot{x} = g(\dot{x}, x)$ is nonconservative. So, strictly speaking, this confirms the fact that a Lagrangian

description, in the sense of §2.24, with $\mathcal{L} \in C^\infty(\mathcal{R}^2)$ can only be found for conservative systems.

3

Systems with Many Degrees of Freedom. Theory of the constraints. Analytical Mechanics

3.1 Systems of Points

We begin with some definitions which are perhaps obvious from the considerations of Chapters 1 and 2, but are nevertheless necessary.

A notational convention that will allow important formal simplifications is that, if $M = m_1 + m_2 + \ldots + m_p$ is a sum of p positive integers, the space $\mathcal{R}^M$ will be considered identical with the space $\mathcal{R}^{m_1} \times \mathcal{R}^{m_2} \times \ldots \times \mathcal{R}^{m_p}$. A point $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_M) \in \mathcal{R}^M$ will be identified with the p-tuple of vectors $(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(p)})$, where $\boldsymbol{\xi}^{(i)} = \big(\xi_{m_1+\ldots+m_{i-1}+1}, \ldots, \xi_{m+1+\ldots+m_i}\big)$ for $i = 1, 2, \ldots, p$.

Very often such a decomposition of $\boldsymbol{\xi}$ into $(Bx^{(1)}, \ldots, \boldsymbol{\xi}^{(p)})$ will be "natural" in the context of the discussion. For instance, if a point in $\mathcal{R}^{3N}$ represents a configuration of a system of N point masses, it will be "natural" to think of $\boldsymbol{\xi}$ as $(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)})$, where $\boldsymbol{\xi}^{(i)} \in \mathcal{R}^3$, $i = 1, \ldots, N$, represents the position in $\mathcal{R}^3$ of the i-th point mass. Every time that it will appear useful, when a natural decomposition of $\boldsymbol{\xi} \in \mathcal{R}^M$ into $(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(p)})$, $\boldsymbol{\xi}^{(i)} \in \mathcal{R}^{m_i}$, $i = 1, \ldots, p$, emerges from the context, $\boldsymbol{\xi}$ will be regarded as a p-tuple of vectors in $\mathcal{R}^{m_1} \times \ldots \times \mathcal{R}^{m_p}$.

Such an identification will be made without explicit mention, provided no real ambiguities arise. Thus, a $\mathcal{R}^{3N}$-valued function $t \to \boldsymbol{\varphi}(t)$ defined on $\mathcal{R}$ will be written, if this is natural within the context, as $t \to (\boldsymbol{\varphi}^{(1)}(t), \ldots, \boldsymbol{\varphi}^{(N)}(t))$ with $t \to \boldsymbol{\varphi}^{(i)}(t)$, $i = 1, \ldots, N$, an $\mathcal{R}^3$-valued function, etc.

If $\mathbf{F}$ is an $\mathcal{R}^d$-valued C^∞-function on $\mathcal{R}^M = \mathcal{R}^{m_1} \times \ldots \times \mathcal{R}^{m_p}$, the Jacobian matrix

$$\Big(\frac{\partial F^{(i)}}{\partial \xi_j}(\boldsymbol{\xi})\Big)_{\substack{i=1,\ldots,d\\ j=1,\ldots,M}}$$

with the symbol $(\partial \mathbf{F}/\partial \boldsymbol{\xi})(\boldsymbol{\xi})$ or $\frac{\partial \mathbf{F}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}$. If $\boldsymbol{\xi} = (\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(p)}) \in \mathcal{R}^{m_1} \times \ldots \times \mathcal{R}^{m_p}$, the symbol $\partial \mathbf{F}/\partial \boldsymbol{\xi}^{(s)}$ will denote the Jacobian matrix $(\partial F^{(i)}/\partial \xi_\ell)(\boldsymbol{\xi})$ where $i = 1, \ldots, d$ and ℓ varies in the set of indices corresponding to the coordinates of $\boldsymbol{\xi}^{(s)}$ (i.e., $\ell = m_1 + \ldots + m_{s-1} + 1, \ldots, m + 1 + \ldots + m_s$).

We can now set up the following definition.

1 Definition. *A "motion" of a system of N point masses in $\mathcal{R}^d$, observed as the time varies in the interval I, is a C^∞ function $t \to \mathbf{x}(t) = (\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(N)}(t))$ defined for $\in I$ and taking values in $\mathcal{R}^{Nd} = \mathcal{R}^d \times \ldots \times \mathcal{R}^d$. A motion $\mathbf{x}$ of a system of N points, with respective masses $m_1, \ldots, m_N > 0$, will be said "governed by a force law $\mathbf{F}$" or "developing under the influence" of a force law $\mathbf{F}$ if:*
(i) $\mathbf{F} = (\mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(N)})$ with $\mathbf{f}^{(i)}$ an $\mathcal{R}^d$-valued function in $C^\infty(\mathcal{R}^{2Nd+1})$, $\forall\, i$.
(ii) For $i = 1, \ldots, N$, $t \in I$:

$$m_i\, \ddot{\mathbf{x}}^{(i)}(t) = \mathbf{f}^{(i)}(\dot{\mathbf{x}}^{(1)}(t), \ldots, \dot{\mathbf{x}}^{(N)}(t), \mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(N)}(t), t) \tag{3.1.1}$$

(iii) Eq. (3.1.1), thought of as a differential equation, is normal for all values of $m_1, \ldots, m_N > 0$ (see Definition 3, §2.5).

Observation. Requirement (iii) is a restriction of "physical nature" on the force laws $\mathbf{F}$ that will be considered. Such laws will often be subject to other restrictions and, always (beginning with the next section), to the condition of verifying the third principle of dynamics (see Chapter 1, §1.3).

A particularly important role will be played by the "conservative force laws", which deserve a formal definition and the rest of the section.

2 Definition. *A force law for a system of N points in $\mathcal{R}^d$, i.e., a function $\mathbf{F} \in C^\infty(\mathcal{R}^{2dN+1})$ with values in $\mathcal{R}^d$, verifying (i) and (iii) of Definition 1, is called "conservative" if:*
(i) it depends solely on the configuration of the system, i.e., there exist N $\mathcal{R}^d$-valued C^∞ functions defined on $\mathcal{R}^{dN}$, $\mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(N)}$, such that

$$\mathbf{f}^{(i)}(\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N)}, \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}, t) \equiv \widetilde{\mathbf{f}}^{(i)}(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}); \tag{3.1.2}$$

(ii) there is a real-valued function $V \in C^\infty(\mathcal{R}^{dN})$ such that for $i = 1, \ldots, N$:

$$\widetilde{\mathbf{f}}^{(i)}(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}) = -\frac{\partial V(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)})}{\partial \boldsymbol{\xi}^{(i)}} \tag{3.1.3}$$

which will be called the "potential energy" of the force law $\mathbf{F}$.

The interest of this definition lies in the fact that the majority of force models are described by conservative force laws, i.e., by force laws that can

be expressed as in Eq. (3.1.3) which, according to the conventions set up at the beginning of this section, means:

$$\widetilde{\mathbf{f}}^{(i)}(\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)})_j = -\frac{\partial V(\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)})}{\partial(\boldsymbol{\xi}^{(i)})_j}. \tag{3.1.4}$$

Furthermore, the energy conservation theorem can be easily extended to systems of N points subject to conservative forces. Given a motion $\mathbf{x}$ of a system of N points, with respective mass $m_1,\ldots,m_N > 0$, define the "kinetic energy" at time t as the quantity

$$T(t) \stackrel{def}{=} \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\mathbf{x}}^{(i)}(t)^2, \tag{3.1.5}$$

while the "potential energy" at time t of the force $\mathbf{F}$ governing the motion, supposed conservative with potential energy function $V \in C^\infty(\mathcal{R}^{dN})$, will be defined as

$$V(t) \stackrel{def}{=} V(\boldsymbol{\xi}^{(1)}(t),\ldots,\boldsymbol{\xi}^{(N)}(t)) \tag{3.1.6}$$

One then notes that

$$\frac{d}{dt}T(t) = \sum_{=1}^{N} m_i\, \dot{\mathbf{x}}^{(i)}(t)\cdot\ddot{\mathbf{x}}^{(i)}(t), \tag{3.1.7}$$

$$\frac{d}{dt}V(t) = \sum_{=1}^{N} \frac{\partial V}{\partial \boldsymbol{\xi}^{(i)}}(\mathbf{x}(t))\cdot\dot{\mathbf{x}}^{(i)}(t), \tag{3.1.8}$$

hence, by Eqs. 3.1.1), 3.1.2), and 3.1.3):

$$\frac{d}{dt}\big(T(t)+V(t)\big) = \sum_{=1}^{N} m_i\, \dot{\mathbf{x}}^{(i)}(t)\big(m_i\ddot{\mathbf{x}}^{(i)}(t) + \frac{\partial V}{\partial \boldsymbol{\xi}^{(i)}}(\mathbf{x}(t))\big) = 0 \tag{3.1.9}$$

Therefore the following proposition holds.

1 Proposition. *If $t \to \mathbf{x}(t) = (\mathbf{x}^{(1)}(t),\ldots,\mathbf{x}^{(N)}(t))$, $t \in I$, is the motion of system of N points, governed by a conservative force law with potential energy V, there is a constant E, "total energy" of the motion, equal at all times to the sum of the kinetic energy and the potential energy:*

$$T(t)+V(t) = E, \quad \forall\, t \in I \tag{3.1.10}$$

with $T(t)$ and $V(t)$ defined in Eqs. (3.1.5) and (3.1.6).

Observation. It is worth stressing that here we are meeting a first but very important difference between one-dimensional and multi-dimensional motions: in the case of the motion of a single point in one dimension, every purely

positional force law is conservative. If $d > 1$ or $N > 1$, there are purely positional force laws which are not conservative in the above sense. For instance, if $N = 1$, $d = 2$, the force law $f_1(\xi_1, \xi_2) = 0$, $f_2(\xi_1, \xi_2) = \xi_1$ is not conservative since $\partial f_1/\partial\xi_2 \neq \partial f_2/\partial\xi_1$, while the two derivatives should coincide if f were conservative (since they would be the mixed second-order derivatives of the same function V).

1. Let f be a $C^\infty(\mathcal{R}^3/\{\mathbf{0}\})$ function with values in $\mathcal{R}^3$ having the form $\mathbf{f}(\mathbf{x}) = \varphi(|\mathbf{x}|)\,\frac{\mathbf{x}}{|\mathbf{x}|}$, $\varphi \in C^\infty(\mathcal{R}_+/\{0\})$. Consider the force law for a system of N point masses given by

$$\mathbf{f}^{(i)}(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}) = \sum_{j \neq i} \varphi(|\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}|)\frac{\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}}{|\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}|} \equiv \sum_{j \neq i} \mathbf{f}(\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}).$$

This force law is defined for configurations such that $\boldsymbol{\xi}^{(i)} \neq \boldsymbol{\xi}^{(j)}$, $\forall\, i \neq j$, and strictly speaking is, therefore, a generalization of the force law notion of Definitions 1 and 2 (requiring the force to be defined for *every* configuration $(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)})$). It will be called conservative if there is a function V, of class C^∞ on the configurations with $\boldsymbol{\xi}^{(i)} \neq \boldsymbol{\xi}^{(j)}$, such that Eq. (3.1.4) holds. In this extended sense, show that the above force law is conservative and

$$V(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}) = \sum_{j < j'} \Phi(|\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}|),$$

where $r \to \Phi(r)$, $r > 0$, is a primitive function to φ: $\Phi(r) = \int^r \varphi(r')\, dr'$.
Find sufficient conditions on φ so that the above force law can be extended by continuity to all configurations becoming a conservative force law in the sense of Definition 2.

2. Let $\Phi_{j,j'}(r)$, $j, j' = 1, \ldots, N$, $j < j'$, be $N(N-1)$ functions in $C^\infty(0, +\infty)$. Consider the force law with potential energy function

$$V(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}) = \sum_{j < j'} \Phi_{j,j'}(|\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}|)$$

Find sufficient conditions on Φ so that the force law is of class $C^\infty(\mathcal{R}^{3N})$.

3.2 Work. Linear and Angular Momentum

One can wonder whether it is possible to extend the energy conservation theorem so that it could be applied to systems subject to nonconservative force laws. The answer is, in some sense, affirmative and it is known as the "alive forces theorem". To formulate this simple theorem, one needs the notion of "work of a force" on a given motion.

3 Definition. *(i) A $\mathcal{R}^{dN}$-valued $C^\infty(\mathcal{R}^{dN+1})$ function $\mathbf{F}$ verifying properties (i) and (iii) of Definition 1 will be called a "force law" for a system of N point masses.*
(ii) If $\mathbf{x}$ is a motion, defined for $t \in I$, of a system of N point masses and if $\boldsymbol{\Phi}$ is a force law for it, not necessarily coinciding with the force law generating the motion $\mathbf{x}$ [i.e. not necessarily verifying Eq. (3.1.1)], one defines the "work" of the force $\boldsymbol{\Phi}$ in the time interval $[t_1, t_2] \subset I$ as the quantity

$$L_{t_1,t_2}(\boldsymbol{\Phi},\mathbf{x}) \stackrel{def}{=} \sum_{i=1}^{N} \int_{t_1}^{t_2} \boldsymbol{\varphi}^{(i)}(\dot{\mathbf{x}}(t),\mathbf{x}(t),t)\,dt \tag{3.2.1}$$

where, following the conventions of §3.1, we set $\boldsymbol{\Phi} = (\boldsymbol{\varphi}^{(1)},\ldots,\boldsymbol{\varphi}^{(N)})$*.*

Observations.

(1) Let $\boldsymbol{\Phi}$ be a purely positional force law, i.e., $\forall\,(\boldsymbol{\eta},\boldsymbol{\xi},t) \in \mathcal{R}^{Nd+1}$:

$$\boldsymbol{\varphi}^{(i)}(\boldsymbol{\eta}^{(1)},\ldots,\boldsymbol{\eta}^{(N)},\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)},t) = \widetilde{\boldsymbol{\varphi}}^{(i)}(\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)}) \tag{3.2.2}$$

$i = 1,\ldots,N$. Then

$$L_{t_1,t_2}(\boldsymbol{\Phi},\mathbf{x}) = \sum_{i=1}^{N} \int_{t_1}^{t_2} \widetilde{\boldsymbol{\varphi}}^{(i)}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}^{(i)}(t)\,dt \tag{3.2.3}$$

and one recognizes in the above integral a line integral of the differential form

$$\sum_{i=1}^{N} \widetilde{\boldsymbol{\varphi}}^{(i)}(\boldsymbol{\xi}) \cdot d\boldsymbol{\xi}^{(i)} \tag{3.2.4}$$

on the curve $\mathcal{I}(\mathbf{x})$ described in $\mathcal{R}^{dN}$ by the point $\mathbf{x}(t)$ as t varies in $[t_1,t_2]$ ("trajectory of $\mathbf{x}$"). Formula (3.2.4) is usually read by saying that the work done by a force on a point which undergoes a displacement is the "scalar product of the force times the displacement".

(2) From observation (1), it follows that the work done by a purely positional force law $\boldsymbol{\varphi}$ in a given time interval during which the system is displaced from the configuration $\mathbf{x}(t_1)$ to $\mathbf{x}(t_2)$ along a certain trajectory $\mathcal{I}$ solely depends upon the trajectory and does not depend on the time law governing the motion along $\mathcal{I}$.

(3) If $\boldsymbol{\Phi}$ is a conservative force with potential energy V [see Eq. 3.1.3)], the differential form of Eq. 3.2.4) coincides with the differential of $-V$:

$$\sum_{i=1}^{N} \widetilde{\boldsymbol{\varphi}}^{(i)}(\boldsymbol{\xi}) \cdot d\boldsymbol{\xi}^{(i)} = -\sum_{i=1}^{N} \frac{\partial V(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}^{(i)}} \cdot d\boldsymbol{\xi}^{(i)} = -dV. \tag{3.2.5}$$

hence, from Eq. (3.2.3), it follows that

$$L_{t_1,t_2}(\boldsymbol{\Phi},\mathbf{x}) = -V(\mathbf{x}(t_2)) + V(\mathbf{x}(t_1)), \tag{3.2.6}$$

showing that the work performed in a given time interval by a conservative force on a motion $\mathbf{x}$ depends solely on the initial and final configurations of the motion, i.e., it is also independent of the trajectory followed by the motion.

The "theorem of the alive forces" can now be formulated.

2 Proposition. *Let* $t \to \mathbf{x}(t)$*,* $t \in I$*, be a motion of a system of* N *points, with masses* $m_1,\ldots,m_N > 0$*, developing in* $\mathcal{R}^d$ *under the action of a force*

law $\mathbf{F}$.
Then the variation of the kinetic energy, or "alive force",[1] *between the times* $t_1, t_2 \in I$, *is equal to the work performed in* $[t_1, t_2]$ *by* $\mathbf{F}$ *on the motion* $\mathbf{x}$*:*

$$T(t_2) - T(t_1) = L_{t_1,t_2}(\mathbf{F}, \mathbf{x}) \tag{3.2.7}$$

Observation. By Eq. (3.2.6), Eq. (3.2.7) becomes the already discussed energy conservation theorem, Proposition 1, whenever $\mathbf{F}$ is conservative.

PROOF. By Definition 1, p.142, of motion developing under the action of a force $\mathbf{F} = (\mathbf{f}^{(1)}, \ldots, \mathbf{f}^{(N)})$, we have

$$m_j\, \ddot{\mathbf{x}}^{(j)} = \mathbf{f}^{(j)}(\dot{\mathbf{x}}^{(1)}, \ldots, \dot{\mathbf{x}}(N), \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}, t), \tag{3.2.8}$$

$\forall\, j = 1, \ldots, N$. Multiplying both sides scalarly by $\dot{\mathbf{x}}^{(j)}$ and summing over j:

$$\sum_{j=1}^{N} m_j \ddot{\mathbf{x}}^{(j)} \cdot \dot{\mathbf{x}}^{(j)} = \sum_{j=1}^{N} \mathbf{f}^{(j)} \cdot \mathbf{x}^{(j)}, \tag{3.2.9}$$

and integrating both sides with respect to t between t_1 and t_2 one finds Eq. (3.2.7). mbe

The interest of Proposition 2 lies in its generality as a consequence of Eq. (3.1.1). There are other immediate consequences of Eq. (3.1.1) valid under the additional assumption that the force law $\mathbf{F}$ governing the motion verifies the third principle of dynamics: they are the so called "cardinal equations" of dynamics, whose interest is also due to their great generality.

As discussed in Chapter 1, the hypothesis that a force law $\mathbf{F}$ for a system of N point masses verifies the third law of dynamics means several things mathematically. First, if $\mathbf{F} = (\mathbf{f}^1, \ldots, \mathbf{f}^{(N)})$, the function $\mathbf{f}^{(j)}$, $j = 1, \ldots, N$, can be represented as

$$\begin{aligned} &\mathbf{f}^{(j)}(\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N)}, \boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}, t) \\ &= \boldsymbol{\varphi}^{(j)e}(\boldsymbol{\eta}^{(j)}, \boldsymbol{\xi}^{(j)}, t) + \sum_{\substack{i=1 \\ i \neq j}} \mathbf{f}^{(i\to j)}(\boldsymbol{\eta}^{(i)}, \boldsymbol{\eta}^{(j)}, \boldsymbol{\xi}^{(i)}, \boldsymbol{\xi}^{(j)}, t) \end{aligned} \tag{3.2.10}$$

where $\mathbf{f}^{(j)} \in C^\infty(\mathcal{R}^{2d+1}), \mathbf{f}^{(i\to j)} \in C^\infty(\mathcal{R}^{4d+1})$ are suitable $\mathcal{R}^d$-valued functions, $\forall\, i, j = 1, \ldots, N$.

For reasons discussed in Chapter 1, the function $\mathbf{f}^{(j)e}$ is called the "external force" acting upon the j-th point mass and $\mathbf{f}^{(i\to j)}$ is called the "force exerted by the i-th point on the j-th one". Second, one assumes that

$$\mathbf{f}^{(i\to j)}(\boldsymbol{\eta}, \boldsymbol{\eta}', \boldsymbol{\xi}, \boldsymbol{\xi}', t) = -\mathbf{f}^{(j\to i)}(\boldsymbol{\eta}', \boldsymbol{\eta}, \boldsymbol{\xi}', \boldsymbol{\xi}, t)) \tag{3.2.11}$$

[1] In the ancient times the alive force was actually defined to be twice the kinetic energy.

and, finally,

$$\mathbf{f}^{(i\to j)}(\boldsymbol{\eta},\boldsymbol{\eta}',\boldsymbol{\xi},\boldsymbol{\xi}',t) \quad \text{is parallel to} \quad \boldsymbol{\xi}'-\boldsymbol{\xi} \tag{3.2.12}$$

Equations (3.2.10)-(3.2.12) are the analytic form taken, in our notations, by the third principle of dynamics for the force law $\mathbf{F}$ acting on the system of point masses under consideration (see, also, Chapter 1).

4 Definition. *A force law for a system of N point masses in $\mathcal{R}^d$ verifies the third principle of dynamics if it admits a representation like Eq. (3.2.10) verifying Eqs. (3.2.11) and (3.2.12). In this case, the quantity*

$$\mathcal{R}^{(e)}(\boldsymbol{\eta}^{(1)},\ldots,\boldsymbol{\eta}^{(N)},\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)},t)=\sum_{i=1}^{N}\mathbf{f}^{(j)e}(\boldsymbol{\eta}^{(j)},\boldsymbol{\xi}^{(j)},t) \tag{3.2.13}$$

thought of as an $\mathcal{R}^d$-valued $C^\infty(\mathcal{R}^{2dN+1})$ function takes the name of "total external force" of the force law $\mathbf{F}$. If $d=3$, the quantity

$$\mathbf{M}^{(e)}_{\boldsymbol{\alpha}}(\boldsymbol{\eta}^{(1)},\ldots,\boldsymbol{\eta}^{(N)},\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)},t)=\sum_{j=1}^{N}(\boldsymbol{\xi}^{(j)}-\boldsymbol{\alpha})\wedge\mathbf{f}^{(j)e}(\boldsymbol{\eta}^{(j)},\boldsymbol{\xi}^{(j)},t)) \tag{3.2.14}$$

is called the "total momentum of the external forces" of $\mathbf{F}$ with respect to the point $\boldsymbol{\alpha}\in\mathcal{R}^d$.

Observation. If $d\neq 3$, it is still possible to define the momentum of the forces with respect to a point: however, it cannot be naturally thought of as a vector in $\mathcal{R}^d$. To avoid complications, rather than on the shaky grounds that the "physical case" is $d=3$, we do not deal with this question.

The following proposition gives the so called "cardinal equations of dynamics":

3 Proposition. *Given a motion $t\to\mathbf{x}(t)$, $t\in\mathcal{R}_+$, of N points in $\mathcal{R}^3$, with masses $m_1,\ldots,m_N>0$, define the "linear momentum" at time t and the "angular momentum", with respect to $\boldsymbol{\alpha}\in\mathcal{R}^3$, at time t as the quantities*

$$\mathbf{Q}(t)\stackrel{def}{=}\sum_{j=1}^{N}m_j\,\dot{\mathbf{x}}^{(j)},\qquad \mathbf{K}_{\boldsymbol{\alpha}}\stackrel{def}{=}\sum_{j=1}^{N}m_j\,(\boldsymbol{\xi}^{(j)}-\boldsymbol{\alpha})\wedge\dot{\mathbf{x}}^{(j)}(t) \tag{3.2.15}$$

If the motion develops under the action of a force law $\mathbf{F}$ verifying the third principle of dynamics and if one shortens $\mathbf{R}^{(e)}(\dot{\mathbf{x}}^{(1)}(t),\ldots,\mathbf{x}^{(N)}(t),t)$ as $\mathbf{R}^{(e)}(t)$ and, likewise, $\mathbf{M}^{(e)}_{\boldsymbol{\alpha}}(\dot{\mathbf{x}}^{(1)}(t),\ldots,\mathbf{x}^{(N)}(t),t)$ as $\mathbf{M}^{(e)}_{\boldsymbol{\alpha}}(t)$ and [see Eqs. (3.2.13),(3.2.14)], then

$$\frac{d}{dt}\mathbf{Q}(t)=\mathbf{R}^{(e)}(t),\qquad \frac{d}{dt}\mathbf{K}_{\boldsymbol{\alpha}}(t)=\mathbf{M}^{(e)}_{\boldsymbol{\alpha}}(t), \tag{3.2.16}$$

Observations.
(1) Sometimes the linear momentum is called the "quantity of motion", while the angular momentum is called "momentum of the quantity of motion". The cardinal equations (3.2.16) show that their time variation depends only upon the external forces acting on the system.
(2) The first cardinal equation in (3.2.16) is often called the "baricenter theorem" or the "center of mass theorem". To understand the origin of this name associate with the motion in $\mathcal{R}^{3N}$, $t \to \mathbf{x}(t) = (\dot{\mathbf{x}}^{(1)}(t), \ldots, \dot{\mathbf{x}}^{(N)}(t))$, $t \in I$, the motion in $\mathcal{R}^3$, $t \to \mathbf{x}_G(t)$, where

$$\mathbf{x}_G(t) = \frac{\sum_{i=1}^{N} m_i \, \mathbf{x}^{(i)}(t)}{\sum_{i=1}^{N} m_i}. \tag{3.2.17}$$

The point $G \stackrel{def}{=} \frac{\sum_{i=1}^{N} m_i \, \mathbf{X}^{(i)}}{\sum_{i=1}^{N} m_i}$ is called the "baricenter" and the motion $t \to \mathbf{x}_G(t)$, $t \in I$, the "baricenter motion". Setting $M = \sum_{i=1}^{N} m_i$ ("total mass of the system"), the first relations in Eqs. (3.2.15) and 3.2.16) become, respectively,

$$\mathbf{Q}(t) = M \dot{\mathbf{x}}_G(t) \tag{3.2.18}$$

$$M \ddot{\mathbf{x}}_G(t) = \mathbf{R}^{(e)}(t) \tag{3.2.19}$$

and Eq. (3.2.19) can be read as "the baricenter of a system of N masses moves as if it were a single point mass subject to the action of a force equal to the total external force acting on the system".

If the external force has the form $\mathbf{f}^{(i)} = m_i \, \mathbf{g} \in \mathcal{R}^3$, "gravity force," the point G has many other nice properties which motivate its name: they are discussed below.
(3) Note that, in general, Eq. (3.2.19) is not a "closed equation": the right-hand side cannot, in fact, be computed without already knowing the locations and the speeds of all the particles of the system. Nevertheless, there are some exceptional particular cases of special importance. For instance, if the external force acting on the j-th point is independent of its position and velocity: *this is the case of the gravity force.*
(4) It is worth stressing that, in general, it is not true that the momentum of the external forces can be computed by imagining the total force as applied to the baricenter; i.e., as $(\mathbf{x}_G - \boldsymbol{\alpha}) \wedge \mathbf{R}^{(e)}$. Neither is it generally true that the derivative of the angular momentum of the baricenter, i.e., of $M(\mathbf{x}_G - \boldsymbol{\alpha}) \wedge \dot{\mathbf{x}}_G$, is the momentum of the total external forces.
(5) However, in the special case

$$\mathbf{f}^{(j)e} = m_i \, \mathbf{g}, \tag{3.2.20}$$

where $\mathbf{g}$ is a fixed vector ("gravity force"), one finds

$$\mathbf{R}^{(e)} = \big(\sum_{i=1}^{N} m_i\big)\,\mathbf{g} \,=\, M\mathbf{g} \tag{3.2.21}$$

and by Eq. (3.2.17),

$$\begin{aligned}\mathbf{M}^{(e)} =& \sum_{j=1}^{N}(\mathbf{x}^{(j)} - \boldsymbol{\alpha}) \wedge m_j\mathbf{g} = \big(\sum_{j=1}^{N} m_j\,(\mathbf{x}^{(j)} - \boldsymbol{\alpha})\big) \wedge \mathbf{g} \\ =& M\,(\mathbf{x}_G - \boldsymbol{\alpha}) \wedge \mathbf{g} = (\mathbf{x}_G - \boldsymbol{\alpha}) \wedge \mathbf{R}^{(e)} = \mathbf{M}^{(e)}_{\boldsymbol{\alpha}}\end{aligned} \tag{3.2.22}$$

Furthermore,

$$\begin{aligned}&\frac{d}{dt}(\mathbf{x}_G - \boldsymbol{\alpha}) \wedge M\dot{\mathbf{x}}_G = (\mathbf{x}_G - \boldsymbol{\alpha}) \wedge M\ddot{\mathbf{x}}_G + \dot{\mathbf{x}}_G \wedge M\dot{\mathbf{x}}_G \\ &\equiv (\mathbf{x}_G - \boldsymbol{\alpha}) \wedge M\ddot{\mathbf{x}}_G = (\mathbf{x}_G - \boldsymbol{\alpha}) \wedge \mathbf{R}^{(e)} = \mathbf{M}^{(e)}_{\boldsymbol{\alpha}};\end{aligned} \tag{3.2.23}$$

i.e., in the case of the gravity forces, the most daring thoughts are allowed: Eqs. (3.2.22) and (3.2.23) show the uniqueness of the gravity force case with respect to the cardinal equations and they explain why the point defined in Eq. (3.2.17) is given the name of “center of gravity”, or “center of mass” or “baricenter”.

Proof. From Eq. (3.2.8), by summing both sides over j, it follows that

$$\sum_{j=1}^{N} m_j\,\ddot{\mathbf{x}}^{(j)} = \sum_{j=1}^{N} \mathbf{f}^{(j)} \tag{3.2.24}$$

but Eqs. (3.2.10) and (3.2.11) and the first of Eqs. (3.2.15) imply $\sum_{j=1}^{N} m_j\,\ddot{\mathbf{x}}^{(j)} = \mathbf{R}^{(e)}$, i.e., the first of Eqs. (3.2.16). Similarly, by externally multiplying both sides of Eq. (3.2.8) by $(\mathbf{x}^{(j)}(t) - \boldsymbol{\alpha})$, $\boldsymbol{\alpha} \in \mathcal{R}^3$, and summing:

$$\begin{aligned}&\sum_{j=1}^{N} m_j\,(\mathbf{x}^{(j)}(t) - \boldsymbol{\alpha}) \wedge \ddot{\mathbf{x}}^{(j)}(t) \\ &= \sum_{j=1}^{N}(\mathbf{x}^{(j)}(t) - \boldsymbol{\alpha}) \wedge \mathbf{f}^{(j)} = \sum_{j=1}^{N}(\mathbf{x}^{(j)}(t) - \boldsymbol{\alpha}) \wedge \mathbf{f}^{(j)e} = \mathbf{M}^{(e)}_{\boldsymbol{\alpha}}\end{aligned} \tag{3.2.25}$$

having used Eqs. (3.2.10) and (3.2.11) and, particularly, Eq. (3.2.12) in the third step to eliminate the contribution of the internal forces:

$$\sum_{j=1}^{N}(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\sum_{\substack{i=1\\ i\neq j}}^{N}\mathbf{f}^{(i\to j)}=\sum_{i\neq j}(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\mathbf{f}^{(i\to j)}$$
$$=\frac{1}{2}\sum_{i\neq j}\Big\{(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\mathbf{f}^{(i\to j)}+(\mathbf{x}^{(i)}(t)-\boldsymbol{\alpha})\wedge\mathbf{f}^{(j\to i)}\Big\} \tag{3.2.26}$$
$$=\frac{1}{2}\sum_{i\neq j}\Big\{(\mathbf{x}^{(j)}(t)-\mathbf{x}^{(i)}(t))\wedge\mathbf{f}^{(i\to j)}\Big\}=\mathbf{0}$$

because $\mathbf{f}^{(i\to j)}=-\mathbf{f}^{(j\to i)}$ and $\mathbf{f}^{(i\to j)}$ is parallel to $\mathbf{x}^{(i)}(t)-\mathbf{x}^{(j)}(t)$, by the third principle. Furthermore,

$$(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\ddot{\mathbf{x}}^{(j)}(t)\equiv(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\frac{d}{dt}\dot{\mathbf{x}}^{(j)}(t)$$
$$=\frac{d}{dt}\big\{(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\dot{\mathbf{x}}^{(j)}(t)\big\}-\big\{\frac{d}{dt}(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\big\}\wedge\dot{\mathbf{x}}^{(j)}(t)$$
$$=\frac{d}{dt}\big\{(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\dot{\mathbf{x}}^{(j)}(t)\big\}-\dot{\mathbf{x}}^{(j)}(t)\wedge\dot{\mathbf{x}}^{(j)}(t) \tag{3.2.27}$$
$$=\frac{d}{dt}\big\{(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\dot{\mathbf{x}}^{(j)}(t)\big\}. \qquad \text{Hence,}$$

$$\sum_{j=1}^{N}m_j(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\ddot{\mathbf{x}}^{(j)}(t)=\frac{d}{dt}\sum_{j=1}^{N}(\mathbf{x}^{(j)}(t)-\boldsymbol{\alpha})\wedge\dot{\mathbf{x}}^{(j)}(t)=\frac{d}{dt}\mathbf{K}_{\boldsymbol{\alpha}} \tag{3.2.28}$$

which, together with Eq. (3.2.25), proves the second equation in (3.2.16).

mbe

1. In Appendix P, there is a table of the masses of the nine main planets and of their distance from the Sun. The mass and radius of the Sun can also be found there. Find the configuration of the planets in which the center of mass of the above ten heavenly bodies is farthest from the center of the Sun and compute the ratio of this distance to the Sun radius. (Assume that the planets move in circular orbits around the Sun.)

2. Same as Problem 1, not counting the Sun.

3. From the data in Appendix P, find the position of the Earth-Moon center of mass relative to the Earth and compare its distance from the center of the Earth with the Earth radius. (Assume the distance between the Earth and Moon to be equal to the maximal or to the minimal distance.)

4. Find the value of the angular momentum of the Earth-Moon system with respect to the center of the Sun, assuming that the latter is fixed in a reference frame with axes fixed with the fixed stars. Assume also that the configuration Moon-Earth-Sun is that of a full lunar eclipse and neglect the orbital inclination of the Moon. Should the angular momentum be time independent? If not, indicate what should be neglected to make it time independent. (*Hint:* The attraction of the Sun on the Earth and on the Moon has vanishing momentum with respect to the center of the Sun, while the Sun-Moon forces are internal forces to the Earth-Moon system.)

5. If $V\in C^\infty(\mathcal{R}^{Nd})$ is bounded below the force law $\mathbf{F}=(\mathbf{f}^{(1)},\ldots,\mathbf{f}^{(N)})$ with $\mathbf{f}^{(i)}\stackrel{def}{=}-\frac{\partial V(\boldsymbol{\xi}^{(i)})}{\partial\boldsymbol{\xi}}$, $i=1,\ldots,N$, is actually a force law in the sense of Definition 3, (i). Show the

validity of this statement. (*Hint:* One need only check condition (iii) of Definition 1. This is obtained by finding an a priori estimate in the sense of §2.5, using energy conservation. Proceed along the lines of the analogous one-dimensional case, §2.5, Proposition 6, p.29).

3.3 The Least Action Principle

The least action principle seen in §2.24 can be extended to systems of N points in $\mathcal{R}^d$ subject to conservative forces. Consider the following definition.

5. Definition.
(i) Let $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ *be the set of motions* $t \to \mathbf{x}(t) = (\mathbf{x}^{(1)}(t), \ldots \mathbf{x}^{(N)}(t)) \in C^\infty([t_1,t_2])$ *such that* $\mathbf{x}(t_1) = \boldsymbol{\xi}_1, \mathbf{x}(t_2) = \boldsymbol{\xi}_2$, *with* $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \mathbf{x}(t) \in \mathcal{R}^{Nd}$.

(ii) If $\mathbf{x} \in \mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$, *the* $\mathcal{V}_\mathbf{x}(\mathcal{M})$ *will denote the space of the "variations" of the motion* $\mathbf{x}$ *in* $\mathcal{M}$*: it is the set of the* $\mathcal{R}^{Nd}$*-valued functions* $\mathbf{y} \in C^\infty([t_1,t_2] \times (-1,1))$, $(t,\varepsilon) \to y(t,\varepsilon)$ *such that:*

$$(a) \qquad \mathbf{y}(t,0) \equiv \mathbf{x}(t),\ \forall\, t \in [t_1,t_2] \tag{3.3.1}$$

$$(b) \qquad \mathbf{y}(t_1,\varepsilon) \equiv \boldsymbol{\xi}_1, \quad \mathbf{y}(t_2,\varepsilon) \equiv \boldsymbol{\xi}_2, \quad \forall\, \varepsilon \in (-1,1) \tag{3.3.2}$$

$$(c) \text{ for all } \varepsilon \in (-1,1), \text{ the function } t \to \mathbf{y}_\varepsilon(t) = \mathbf{y}(t,\varepsilon), \quad t \in [t_1,t_2] \tag{3.3.3}$$

is a motion $\mathbf{y}_\varepsilon \in \mathcal{M}$. *We shall set* $\mathcal{V}_\mathbf{x}(\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)) \equiv \mathcal{V}_\mathbf{x}$

(iii) If $\mathcal{L} \in C^\infty(\mathcal{R}^{2Nd+1})$ *is a real-valued function, define the action with Lagrangian density* $\mathcal{L}$ *as the real-valued function* A *on* $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$*:*

$$A(\mathbf{x}) \stackrel{def}{=} \int_{t_1}^{t_2} \mathcal{L}(\dot{\mathbf{x}}(t), \mathbf{x}(t), t)\, dt \tag{3.3.4}$$

(iv) The action A *in Eq. (3.3.4) is said to be stationary or locally minimal the motion* $\mathbf{x} \in \mathcal{M} \subset \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ *if the function* $\varepsilon \to A(\mathbf{y}_\varepsilon)$, $\varepsilon \in (-1,1)$, *is stationary or locally minimal for* $\varepsilon = 0$ *and for all* $\mathbf{y} \in \mathcal{V}_\mathbf{x}(\mathcal{M})$.

The stationarity condition of Eq. (3.3.4) on all of $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ in $\mathbf{x}$ is deduced exactly along the same lines and patterns followed to prove the analogous condition seen in Proposition 36, §2.24, p.130, through the principle of the vanishing integrals (Appendix D). Therefore, the detailed proof of the following proposition is left to the reader.

4 Proposition. *A motion* $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ *is a stationary point in* $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ *for the action of Eq. (3.3.4) if and only if*

$$\frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\eta}^{(i)}}(\dot{\mathbf{x}}(t), \mathbf{x}(t), t)\Big) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\xi}^{(i)}}(\dot{\mathbf{x}}(t), \mathbf{x}(t), t) \tag{3.3.5}$$

for all $t \in [t_1,t_2]$ *and for all* $i = 1, 2, \ldots, N$.

Observations.
(1) In Eq. (3.3.5) we use the notation on the derivatives introduced in §3.1.
(2) Often, with an abuse of notation, Eq. (3.3.5) is compactly written as

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} = \frac{\partial \mathcal{L}}{\partial \mathbf{x}}. \tag{3.3.6}$$

An immediate corollary to Proposition 4 is the following.

5 Proposition. *Given a real-valued $C^\infty(\mathcal{R}^{Nd})$ function V, bounded from below,[2] the motion $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ makes the action with Lagrangian density*

$$\mathcal{L}(\boldsymbol{\eta}, \boldsymbol{\xi}, t) = \frac{1}{2}\sum_{i=1}^{N} m_i \, {\boldsymbol{\eta}^{(i)}}^2 - V(\boldsymbol{\xi}) \tag{3.3.7}$$

$m_j > 0, j = 1, \ldots, N$, *stationary on $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ if and only if $\mathbf{x}$ is a motion of N points in $\mathcal{R}^d$ with masses $m_1, \ldots, m_N > 0$ which, for $t \in [t_1, t_2]$, develops subject to influence of the force law $\mathbf{F}$ with potential energy V.*

Proof. It is enough to substitute Eq. (3.3.7) into Eq. (3.3.5) to see that, in this case, Eq. (3.3.5) becomes Eq. (3.1.1) with $\mathbf{F}$ given by Eq. (3.1.3), i.e.,

$$m_j \, \ddot{\mathbf{x}}^{(j)}(t) = -\frac{\partial V}{\partial \boldsymbol{\xi}^{(j)}}, \qquad j = 1, \ldots, N \tag{3.3.8}$$

if $\mathbf{x}(t) = (\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(N)}(t))$, $t \in [t_1, t_2]$. mbe

The following generalization of Proposition 38, §2.24, is also valid, but the proof is left as a problem for the reader (since it is an essentially a word-by-word repetition of that of Proposition 38, p.132).

6 Proposition. *Let $t \to \mathbf{x}(t)$, $t \in \mathcal{R}_+$, be a motion of a system of N points in $\mathcal{R}^d$, with masses $m_1, \ldots, m_N > 0$, developing under the action of a conservative force with potential energy $V \in C^\infty(\mathcal{R}^{Nd})$. Given $t_1 \in \mathcal{R}_+$ and $t_2 > t_1$, if $t_2 - t_1$, is small enough, the motion $t \to \mathbf{x}(t)$ considered in the time interval $[t_1, t_2]$ is a point of local minimum in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ for the action with Lagrangian (3.3.7).*

The comments seen at the end of Chapter 2, pp.133-135, extend to the contents of this section. It is, in particular, quite important that the reader extends to the case of a system of point masses the observations made in §2.24, concerning the representations of motions in coordinates other than Cartesian coordinates and concerning the invariance of the Lagrange equations (3.3.6) with respect to changes in coordinates (see §2.24, p.133 and following).

In the following sections and in their exercises, we shall see some interesting applications of the "Lagrangian formulation" (3.3.6) of the equations of motion as a "change of coordinates invariant" formulation of such equations. Among these will be the theory of perfect constraints.

[2] See Problem 5, §3.2.

3.4 Introduction to the Constrained Motion Theory

Elli avien cappe con cappucci bassi
Dinanzi a li occhi, fatte della taglia
Che in Clugnì per li monaci fassi.
Di fuor dorate son, si ch'elli abbaglia;
Ma dentro tutte piombo, e gravi tanto ...[3]

The principle of least action inspires the following somewhat trivial considerations. Let $\mathbf{x} \to A(\mathbf{x})$ be the action of Eqs. (3.3.4) and (3.3.7) defined on the motions in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ of a system of N point masses subject to a conservative force $\mathbf{F}$.

Suppose a priori known that the force law is such that the motion $\mathbf{x}$ that develops under its influence from $\boldsymbol{\xi}_1$ to $\boldsymbol{\xi}_2$, within times t_1 and t_2, verifies some properties like $|\mathbf{x}(t)| \le S$ or $|\ddot{\mathbf{x}}(t)| \le P$ or $|\mathbf{x}^{(1)}(t)| = \mathbf{0}$, etc. Then it is clear that the research of $\mathbf{x}$ in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ can be restricted to the subset $\mathcal{M}$, of the motions in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ verifying the properties under consideration.

Very often it happens that a system of point masses is subject lo "constraints", i.e., to force laws that allow only a "few" motions among those a priori possible, at least for vast classes of initial data. Think of a point mass constrained to remain on a surface: in this case, the surface acts on the point with a force systematically such as to forbid the abandonment of the surface itself by the point, whenever the initial data $(\boldsymbol{\eta}, \boldsymbol{\xi})$ have $\boldsymbol{\xi}$ on the surface and $\boldsymbol{\eta}$ tangent to it.

Think, also, of a rigid system of N points. Now the i-th point will exert on the j-th point a force $f^{(i\to j)}$ systematically such that the two points remain at a fixed distance from each other.

By taking into account the constraints, the allowable motions in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ will generally be parameterizable with ℓ coordinates, and often $\ell \ll Nd$; consequently, it will be possible to imagine a description of the motions in terms of ℓ functions of time. Therefore, the Lagrangian and the action will also be expressible in terms of the same ℓ functions, and the action of a motion $\mathbf{x}$ allowed by the constraints will take the form

$$A(\mathbf{x}) = \int_{t_1}^{t_2} dt\, \widetilde{\mathcal{L}}(\dot{a}_1(t), \dots, \dot{a}_\ell(t), a_1(t), \dots, a_\ell(t), t) \tag{3.4.1}$$

if $t \to (a_1(t), \dots, a_\ell(t))$, $t \in [t_1, t_2]$, is the description of the motion $\mathbf{x}$ in the ℓ "essential coordinates".

---

[3] In basic English:
They had capes with low hoods
in front of the eyes, made in the fashion
that in Cluny is used for the monks.
Golden they are outside, so that they dazzle
but inside they are all leaden and heavy a lot ...
(Dante, Inferno, Canto XXIII).

To be less vague, assume that there are N $\mathcal{R}^d$-valued functions in $C^\infty(\mathcal{R}^\ell)$:

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell) \to \mathbf{X}^{(i)}(\boldsymbol{\alpha}) = \mathbf{X}^{(i)}(\alpha_1, \ldots, \alpha_\ell), \tag{3.4.2}$$

$i = 1, \ldots, N$, such that the set of the motions $t \to \mathbf{x}(t) = (\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(N)}(t))$ $t \in [t_1, t_2]$, which are "constrained" or "allowed" by the constraints is simply the set of the motions which is the image of the motions in $\mathcal{R}^\ell$ via the transformation (3.4.2). Thus, given a motion $t \to \mathbf{a}(t)$, $t \in [t_1, t_2]$, in $\mathcal{R}^\ell$ one describes, via Eq. (3.4.2), the constrained motion $t \to \mathbf{x}(t)$, $t \in [t_1, t_2]$, where

$$\mathbf{x}^{(i)}(t) = \mathbf{X}^{(i)}(\mathbf{a}(t)), \qquad i = 1, 2, \ldots, N \tag{3.4.3}$$

which we shorten as $\mathbf{x}(t) = \mathbf{X}(\mathbf{a}(t))$.

In other words, let us admit that the conservative force law $\mathbf{F}$ for the system of N point masses under consideration is such that the motions in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ that can actually develop under its influence starting from a given class of initial data are necessarily contained in the class of the motions having the form of Eq. (3.4.3) with $\mathbf{a} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$, where $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathcal{R}^\ell$ and $\mathbf{X}(\boldsymbol{\alpha}_1) = \boldsymbol{\xi}_1, \mathbf{X}(\boldsymbol{\alpha}_2) = \boldsymbol{\xi}_2$.

If $\mathbf{x}$ is a constrained motion in the sense just discussed, its action, Eq. (3.3.4), with respect to the Lagrangian (3.3.7), where V is the potential energy of $\mathbf{F}$, can be written as in Eq. (3.4.1) if $\mathcal{L} \in C^\infty(\mathcal{R}^{2\ell+1})$ is the function

$$\widetilde{\mathcal{L}}(\beta_1, \ldots, \beta_\ell, \alpha_1, \ldots, \alpha_\ell, t) = \frac{1}{2} \sum_{i=1}^{N} m_i \Big(\sum_{j=1}^{N} \frac{\partial \mathbf{X}^{(i)}}{\partial \alpha_j} \beta_j \Big)^2 - V(\mathbf{X}(\boldsymbol{\alpha})), \tag{3.4.4}$$

because $\dot{\mathbf{x}}^{(i)}(t)$ can be computed, by differentiating Eq. (3.4.3), as

$$\dot{\mathbf{x}}^{(i)}(t) = \sum_{j=1}^{N} \frac{\partial \mathbf{X}^{(i)}}{\partial \alpha_j}(\boldsymbol{\alpha}(t))\, \dot{a}_j(t), \qquad j = 1, 2, \ldots, N, \tag{3.4.5}$$

whenever $\mathbf{x}$ is the constrained motion image of $\mathbf{a} : \mathbf{x} = \mathbf{X}(\mathbf{a})$.

Hence, if $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ is the motion that actually develops under the influence of the force F and if $\mathbf{x}$ is the image via Eq. (3.4.3) of $\mathbf{a}$, then the action A with Lagrangian (3.3.7) is stationary in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ on $\mathbf{x}$, while the action $\widetilde{A}$ with Lagrangian given by Eq. (3.4.4) is stationary on $\mathbf{a}$ in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$. This property is an immediate consequence of the fact that if A is stationary on a motion $\mathbf{x}$ in it is also stationary on $\mathbf{x}$ in any smaller set $\mathcal{M}' \subset \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ provided $\mathbf{x} \in \mathcal{M}'$. In our case, through Eq. (3.4.3), $\mathcal{M}'$ would be the set of the motions which is the image of the motions in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$.

By Proposition 4, §3.3, the stationarity condition for A, i.e., for the action on $\mathcal{M}_{t_1,t_2}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ with Lagrangian density (3.4.4), is

$$\frac{d}{dt}\frac{\partial\widetilde{\mathcal{L}}}{\partial\beta_i}(\dot{\mathbf{a}}(t),\mathbf{a}(t),t)=\frac{\partial\widetilde{\mathcal{L}}}{\partial\alpha_i} \tag{3.4.6}$$

$i=1,2,\ldots,\ell,\ \forall\, t\in[t_1,t_2]$.

The importance of the above considerations is easily realized: Eq. (3.4.6) is already the equation of motion after the elimination of the parameters describing the system, necessary a priori but made "useless" or "redundant" by the presence of the constraints which allow one to reduce the number of the coordinates needed to describe the actually "possible" configurations, from Nd down to ℓ via (3.4.2) and (3.4.3).

Therefore, the idea occurs that the mechanism for the elimination of the redundant coordinates in conservative systems subject to simple constraints, like Eqs. (3.4.2) and (3.4.3), might be particularly simple: it will be enough to rewrite the Lagrangian density of the action only in terms of the essential coordinates through Eq. (3.4.2) and, then, deduce Eq. (3.4.6).

However, the principle of conservation of difficulties makes it clear that there must be some serious obstacle to the actual applications of such a shining but simplistic vision.

The true constraints are, in fact, generated by forces that, as we shall see shortly, generally are neither simple nor conservative (in the sense of Definition 2, p.142, §3.1) but depend on the velocities of the points as well as on their positions.

In such situations, the above considerations become essentially useless since they are not applicable to the simplest and most interesting motions constrained in the sense that they are parameterizable as in Eqs. (3.4.2) and (3.4.3), by ℓ coordinates.

To understand better what has just been said, let us consider the case of a point constrained to stay on a curve $\Gamma\subset\mathcal{R}^3$ with intrinsic parametric equations given by

$$s\to\boldsymbol{\xi}(s),\qquad s\in\mathcal{R} \tag{3.4.7}$$

where s is the curvilinear abscissa on Γ (which will be supposed to be a simple curve, i.e., without double points and open). Assume that the curve Γ exerts a force on the point mass which keeps it on Γ for all motions starting from initial data $(\boldsymbol{\eta},\boldsymbol{\xi})$ with $\boldsymbol{\xi}=\boldsymbol{\xi}(s_0)$, $\boldsymbol{\eta}=\frac{d\boldsymbol{\xi}}{ds}(s_0)\dot{s}_0$ (i.e., with $\boldsymbol{\xi}\in\Gamma$ and $\boldsymbol{\eta}$ tangent to it), with $(s_0,\dot{s}_0)\in\mathcal{R}^2$.

If $\boldsymbol{\tau}(s),\mathbf{n}(s)$ denote, respectively, the tangent and the principal normal versors to Γ at the point with curvilinear abscissa s and if $r(s)$ denotes the curvature radius at the same point, it is well known that

$$\boldsymbol{\tau}(s)=\frac{d\boldsymbol{\xi}(s)}{ds},\quad \frac{\mathbf{n}(s)}{r(s)}=\frac{d\boldsymbol{\tau}(s)}{ds} \tag{3.4.8}$$

Then if $t\to s(t), t\in\mathcal{R}$, is a motion on Γ described by the time variation of the curvilinear abscissa, we find

$$\frac{d}{dt}\boldsymbol{\xi}(s(t)) = \dot{s}(t)\,\boldsymbol{\tau}(s(t)) \tag{3.4.9}$$

and

$$\frac{d^2}{dt^2}\boldsymbol{\xi}(s(t)) = \ddot{s}(t)\,\boldsymbol{\tau}(s(t)) + \frac{\dot{s}(t)^2}{r(s(t))}\,\mathbf{n}(s(t)). \tag{3.4.10}$$

If the point is subject to a force which is the sum of the constraint reaction $\mathbf{R}(\dot{s}, s)$ and of an external force $\mathbf{f}(s)$, then

$$m\,\ddot{\mathbf{x}} = \mathbf{f} + \mathbf{R} \tag{3.4.11}$$

if $m > 0$ is the mass and $\mathbf{x}(t) = \boldsymbol{\xi}(s(t))$ denotes the motion in $\mathcal{R}^3$.

By Eq. (3.4.10), Eq. (3.4.11) becomes

$$m\ddot{s} = \mathbf{f}\cdot\boldsymbol{\tau} + \mathbf{R}\cdot\boldsymbol{\tau}, \qquad m\frac{\dot{s}^2}{r} = \mathbf{f}\cdot\mathbf{n} + \mathbf{R}\cdot\mathbf{n} \tag{3.4.12}$$

and from the second equation, it follows that the normal component of the constraint reaction is

$$\mathbf{R}\cdot\mathbf{n} = m\,\frac{\dot{s}^2}{r(s)} - \mathbf{f}(s)\cdot\mathbf{n}(s) \tag{3.4.13}$$

at the point of Γ with coordinate s when it is occupied by a mass m with speed along Γ given by $\dot{s}$.

From Eq. (3.4.13), one sees that $\mathbf{R}(\dot{s}, s)$ is necessarily $\dot{s}$ dependent if $0 < r(s) < +\infty$, as will be supposed, and therefore the constraint reaction cannot be conservative in the very restrictive sense of §3.1.

Nevertheless, the essence of the idea which arose in connection with Eq. (3.4.6) will be saved: it will, however, be necessary to go through a long analysis which, as is to be expected, involves a deeper physico-mathematical discussion of the notion of constraint. Such a discussion will be aimed at clarifying the definition of constraint, i.e., the physical phenomenon mathematically modeled as a "constraint".

In the next section a general mathematical definition of constraint will be presented, stressing its main mathematical properties and delaying until the later sections a deeper discussion showing how the empirical notion of a frictionless constraint is naturally schematized by the introduced mathematical structures.

3.4.1 Exercises

1. Let Γ be a circle in $\mathcal{R}^3$ with radius r. Find $r(s), \mathbf{n}(s), \boldsymbol{\tau}(s)$ [see Eq. (3.4.8)].

2. Let Γ be an ellipse with equations $z = 0, x^2/a^2 + y^2/b^2 = 1$, $a, b > 0$. Find $r(s), \mathbf{n}(s), \boldsymbol{\tau}(s)$, at the point $(x, y, 0)$.

3. Show that the force law $\mathbf{R}(\dot{\mathbf{x}}, \mathbf{x}) = -\frac{m\dot{\mathbf{x}}^2}{r^2}\,\mathbf{x}$, $(\dot{\mathbf{x}}, \mathbf{x}) \in \mathcal{R}^2 \times \mathcal{R}^2$ produces a constraint for the motions of a point with mass $m > 0$ with initial data $(\boldsymbol{\eta}, \boldsymbol{\xi})$ with $\boldsymbol{\eta}\cdot\boldsymbol{\xi} = 0$, $|\boldsymbol{\xi}| = r$. The

constraint is to the circle $\Gamma = \{\boldsymbol{\xi} \in \mathcal{R}^2,\ |\boldsymbol{\xi}| = r\}$. (*Hint:* Show that the circular uniform motion verifies the equations of motions and use the uniqueness theorem.)

4. Same as Problem 3 in $\mathcal{R}^3$, replacing the circle Γ with the surface of a sphere.

5. Same as Problem 3, using Archimedes' spiral (with equations $\varrho = a\,\theta$, $a > 0$, in polar coordinates), finding an appropriate force $\mathbf{R}$ producing a constraint to the spiral.

6. Find an appropriate force $\mathbf{R}$ producing a constraint to Γ, as defined in Problems 3-5, if the point mass is also subject to a conservative force with potential energy $V = \frac{\kappa}{2}\mathbf{x}^2$, $\kappa > 0$.

7. Show that no purely positional force law $\mathbf{R}$ can force every motion with initial data $(\boldsymbol{\eta}, \boldsymbol{\xi})$ with $\boldsymbol{\eta} \cdot \boldsymbol{\xi} = 0$, $|\boldsymbol{\xi}| = 1$, to move on the unit circle in $\mathcal{R}^2$, regardless of the mass m of the point. (*Hint:* Let $(\boldsymbol{\eta}_0, \boldsymbol{\xi}_0)$ be an initial datum at $t = 0$ producing a motion which stays on the circle. Consider the motion with initial datum $(2\boldsymbol{\eta}_0, \boldsymbol{\xi}_0)$ and show that it must abandon the unit circle, for $t > 0$ and small, by using the Lagrange-Taylor theorem or, alternatively, by using Eq. (3.4.13).)

3.5 Ideal Constraints as Mathematical Entities

The following is a rather general mathematical definition of a constrained motion for a system of N point masses.

6 Definition. *Given s real-valued $C^\infty(\mathcal{R}^{2Nd+1})$-functions $\psi^{(1)}, \ldots, \psi^{(s)}$ we shall say that a system of N points, with masses $m_1 \ldots, m_N > 0$, subject to a force law $\mathbf{F}$ is constrained by the constraints $\psi^{(1)}, \ldots, \psi^{(s)}$ if $\mathbf{F}$ is such that the motions $t \to \mathbf{x}(t) = (\mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(N)}(t))$, $t \in \mathcal{R}_+$, developing under its influence identically verify the s relations, $i = 1, \ldots, s$:*

$$\psi^{(i)}(\dot{\mathbf{x}}^{(1)}(t), \ldots, \dot{\mathbf{x}}^{(N)}(t), \mathbf{x}^{(1)}(t), \ldots, \mathbf{x}^{(N)}(t), t) = 0 \tag{3.5.1}$$

$\forall t \in \mathcal{R}_+$, provided there is a time t (e.g., $t = 0$) when Eq. (3.5.1) holds.

Examples
(1) If $V \in C^\infty(\mathcal{R}^{Nd})$ and $E \in \mathcal{R}$, the function

$$\begin{aligned}&\psi^{(i)}(\boldsymbol{\eta}^{(1)}(t), \ldots, \boldsymbol{\eta}^{(N)}(t), \boldsymbol{\xi}^{(1)}(t), \ldots, \boldsymbol{\xi}^{(N)}(t), t) \\ &= \frac{1}{2}\sum_{i=1}^{N} m_i {\boldsymbol{\eta}^{(i)}}^2 - V(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}) - E\end{aligned} \tag{3.5.2}$$

is a constraint for the motions of a system of N point masses with masses $mi, \ldots, m'' > 0$ subject to a force law with V as a potential energy.
(2) Given a system of N points, with masses $m_1, \ldots, m_N > 0$ subject to a force law $\mathbf{F}$ verifying the third principle of dynamics and with zero external forces, let $\mathbf{q}, \mathbf{m} \in \mathcal{R}^3$. Define the six functions on $\mathcal{R}^{2Nd+1}$ (actually independent on the last coordinate):

$$\psi(\boldsymbol{\eta}^{(1)}(t),\ldots,\boldsymbol{\eta}^{(N)}(t),\boldsymbol{\xi}^{(1)}(t),\ldots,\boldsymbol{\xi}^{(N)}(t),t) = \sum_{i=1}^{N} m_i\,\boldsymbol{\eta}_i - \mathbf{q},$$

$$\psi'(\boldsymbol{\eta}^{(1)}(t),\ldots,\boldsymbol{\eta}^{(N)}(t),\boldsymbol{\xi}^{(1)}(t),\ldots,\boldsymbol{\xi}^{(N)}(t),t) = \sum_{i=1}^{N} m_i\,\boldsymbol{\xi}^{(i)} \wedge \boldsymbol{\eta}^{(i)} - \mathbf{m}, \tag{3.5.3}$$

then the above six functions provide six constraints for the system.
(3) More generally, every conservation law may be interpreted as a constraint.
(4) The above examples may be pushed to the extremes: given $(\boldsymbol{\eta}_0,\boldsymbol{\xi}_0) \in \mathcal{R}^{2Nd}$ and calling S_t, the evolution flow associated with a time independent force law F acting on a system of N point masses, the $2Nd$ functions:

$$\boldsymbol{\psi}(\boldsymbol{\eta},\boldsymbol{\xi},t) = S_t(\boldsymbol{\eta},\boldsymbol{\xi}) - (\boldsymbol{\eta}_0,\boldsymbol{\xi}_0) \tag{3.5.4}$$

are constraints for the system.
(5) Consider a point with mass $m > 0$ in $\mathcal{R}^3$ subject to a force law given by

$$\mathbf{F}(\boldsymbol{\eta},\boldsymbol{\xi}) = -m\frac{\boldsymbol{\eta}^2}{r}\frac{\boldsymbol{\xi}}{r} \tag{3.5.5}$$

where $r > 0$ is constant. Then the following two functions:

$$\psi_1(\boldsymbol{\eta},\boldsymbol{\xi}) = (\boldsymbol{\xi}^2 - r^2)^2 + (\boldsymbol{\eta}\cdot\boldsymbol{\xi})^2, \qquad \psi_2(\boldsymbol{\eta},\boldsymbol{\xi}) = \xi_3^2 + \eta_3^2 \tag{3.5.6}$$

are constraints for the system (see Problem 3, §3.4). are constraints for the system.

Observation. The above examples of constraints may leave the reader a bit perplexed, particularly Example 4. In some sense it shows that all the motions can be considered as constrained motions.
It will be seen that the constraints become interesting only when they can actually be "constructed", so that they can be used to reduce the number of degrees of freedom, or of parameters, necessary to describe the motions. A constraint of the type in the Example 4 is of little use in practice since it can be constructed only when all the motions of the system are perfectly understood (i.e., when S_t is a "known transformation"). However, this is usually the aim of the theory and it cannot be considered as a starting point.

Particularly interesting are the velocity-independent and time-independent constraints.

7 Definition. *In the context of Definition 6, assume that there exist s real-valued functions in $C^\infty(\mathcal{R}^{Nd})$, $\varphi^{(1)},\ldots,\varphi^{(s)}$, such that, $\forall\, i = 1,\ldots,s$,*

$$\psi^{(i)}(\boldsymbol{\eta}^{(1)},\ldots,\boldsymbol{\eta}^{(N)},\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)},t) =\equiv \varphi^{(i)}(\boldsymbol{\xi}^{(1)},\ldots,\boldsymbol{\xi}^{(N)}) \tag{3.5.7}$$

for all $(\boldsymbol{\eta}, \boldsymbol{\xi}, t) \in \mathcal{R}^{2Nd+1}$. *We shall say that the system is "subject to s holonomous constraints* $\varphi^{(1)}, \ldots, \varphi^{(s)}$*".*[4] *We shall denote* $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$ *the subset of the motions in* $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ *consisting of the motions* $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ *such that*

$$\varphi^{(i)}(\mathbf{x}(t)) \equiv 0 \tag{3.5.8}$$

This set will be called the set of the motions "subject to" or "compatible with" the constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$.
Finally, if $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathcal{R}^{Nd}$, *we shall say that* $\boldsymbol{\xi}$ *is a configuration "compatible with the constraints" if* $\varphi^{(i)}(\boldsymbol{\xi}) = 0$, $j = 1, \ldots, s$, *and that* $\boldsymbol{\eta}$ *is a velocity "compatible with the constraints in* $\boldsymbol{\xi}$*" if there is a motion* $t \to \mathbf{x}(t)$*, defined for* t *near zero, such that* $\mathbf{x}(0) = \boldsymbol{\xi}, \dot{\mathbf{x}}(0) = \boldsymbol{\eta}$ *and* $\mathbf{x}(t)$ *is compatible with the constraints for all* t.

Observations.
(1) By the assumed time invariance of the constraints [see Eq. (3.5.7)], the choice of time $t = 0$ in the last part of Definition 7 has no special meaning.
2) In Problem 2 at the end of this section, we mention that when the vectors $\frac{\partial \varphi^{(i)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}$, $j = 1, \ldots, s$, are s linearly independent vectors in $\mathcal{R}^{Nd}$, the constraint compatibility condition for a velocity $\boldsymbol{\eta}$ can be analytically expressed as

$$\frac{\partial \varphi^{(i)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \cdot \boldsymbol{\eta} = 0, \qquad j = 1, \ldots, s \tag{3.5.9}$$

which has a clear geometrical meaning.
(3) Given $\boldsymbol{\xi} \in \mathcal{R}^{Nd}$ compatible with the constraints, the set of the velocity vectors $\boldsymbol{\eta}$ compatible with the constraints in $\boldsymbol{\xi}$ is always nonempty since it contains $\boldsymbol{\eta} = \mathbf{0}$.

Our first task will now be to set up a precise definition of a "perfect holonomous constraint". A possible definition is inspired by Eq. (3.4.12): in that case, the constraint to the line Γ is naturally called "ideal" if $\mathbf{R} \cdot \boldsymbol{\tau} = 0$, i.e., if the "only effect" of the constraint is to keep the motion on Γ; in fact, the equation of motion simply becomes

$$m\, \ddot{s}(t) = \mathbf{f}(s) \cdot \boldsymbol{\tau}(s) \tag{3.5.10}$$

which can be read "the acceleration along Γ is proportional to the projection on Γ of the active force", i.e., of the part of the force distinct from the constraint reaction.

The relation $\mathbf{R} \cdot \boldsymbol{\tau} = 0$ means that the reaction acts orthogonally to Γ. However, it is not immediately clear what should be meant by the reaction being orthogonal to the constraint in the case of the general constraints considered in Definition 7.

[4] Holonomous simply means "depending on the site".

After some thought, the following notion appears natural: the constraint reaction or, more generally, a force law $\mathbf{R}(\boldsymbol{\eta}, \boldsymbol{\xi})$ acting on a system of N points masses in $\mathcal{R}^d$ occupying the configuration $\boldsymbol{\xi}$ with velocity $\boldsymbol{\eta}$ (*both constraint compatible*), is "orthogonal to the constraint" if, calling $\boldsymbol{\eta}'$ the velocity of any other *constraint compatible* motion at the time when it occupies the configuration $\boldsymbol{\xi}$, it is

$$\mathbf{R}(\boldsymbol{\eta}, \boldsymbol{\xi}) \cdot \boldsymbol{\eta}' = 0, \tag{3.5.11}$$

which, more explicitly, is

$$\sum_{j=1}^{N} \boldsymbol{\eta}'^{(j)} \cdot \mathbf{R}^{(j)}(\boldsymbol{\eta}, \boldsymbol{\xi}) = 0 \tag{3.5.12}$$

One could argue and debate about this extension. However, in this section we shall first investigate its mathematical meaning, delaying the discussion of its deep and interesting physical interpretation until later on. Let us therefore establish the following definition.

8 Definition. *Let* $\mathbf{F}$ *be a time-independent force law for a system of* N *points in* $\mathcal{R}^d$*. Assume that* $\mathbf{F}$ *produces* s *holonomous constraints* $\varphi^{(1)}, \ldots, \varphi^{(s)}$*.*
Given a positional force law $\mathbf{F}^{(a)} \in C^\infty(\mathcal{R}^{Nd})$ *for the system, we define the "constraints reaction" with respect to the "active force"* $\mathbf{F}^{(a)}$ *as the quantity* $\mathbf{R} = \mathbf{F} - \mathbf{F}^{(a)}$*. Furthermore, we shall say that the system of constraints is "ideal" with respect to the pair* $(\mathbf{R}, \mathbf{F}^{(a)})$ *if for all* $\boldsymbol{\xi} \in \mathcal{R}^{Nd}$ *compatible with the constraints, i.e., such that* $\varphi^{(j)}(\boldsymbol{\xi}) \equiv 0, \forall\, j$ *(see Definition 7), it is*

$$\mathbf{R}(\boldsymbol{\eta}_1, \boldsymbol{\xi}) \cdot \boldsymbol{\eta}_2 = 0 \tag{3.5.13}$$

for all choices of constraint compatible velocity vectors $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ *(in the sense of Definition 7). We shall refer to this situation by using the shortened locution "the system of point masses is subject to the active force* $\mathbf{F}^{(a)}$ *and to* s *holonomous ideal constraints* $\varphi^{(1)}, \ldots, \varphi^{(s)}$*.*

Observations.
(1) Therefore, the last sentence means that the system is subject to a time-independent force law $\mathbf{F}$ producing the constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$. which are ideal with respect to the active force $\mathbf{F}^{(a)}$ and to the "reaction" $\mathbf{R} = \mathbf{F} - \mathbf{F}^{(a)}$. Strictly speaking, the last sentence of Definition 8 should be subject to a consistency check: in terms of the information contained in it, it should be possible to reconstruct the equations of motion at least as far as the constrained motions are concerned; i.e., given $\mathbf{F}^{(a)}$ and the constraints it should be possible to reconstruct $\mathbf{F}(\boldsymbol{\eta}, \boldsymbol{\xi})$ for all constraint compatible $(\boldsymbol{\eta}, \boldsymbol{\xi})$. This is actually possible and, basically, it is the content of Proposition 8 (below) and of the first observation to it (see, also, Problem 2 at the end of this section).
(2) It is important to stress that the decomposition $\mathbf{F}^{(a)} + \mathbf{R}$ of the force as a sum of an "active force" and of an "ideal constraint reaction" is certainly

not unique, if it exists at all. For instance, if $\mathbf{F}$ is a conservative force field for our system whose potential energy $\widetilde{V} \in C^{\infty}(\mathcal{R}^{Nd})$ is constant on the region of $\mathcal{R}^{Nd}$, where

$$\varphi^{(1)}(\boldsymbol{\xi}) = \ldots = \varphi^{(s)}(\boldsymbol{\xi}) = 0, \tag{3.5.14}$$

then the decomposition

$$\mathbf{F} = (\mathbf{F}^{(a)} + \widetilde{\mathbf{F}}) + (\mathbf{R} - \widetilde{\mathbf{F}}) \tag{3.5.15}$$

can be shown to be another decomposition of $\mathbf{F}$ into an "active" part $\mathbf{F}^{(a)} + \widetilde{\mathbf{F}}$ and into a "reaction" $\mathbf{R}' = \mathbf{R} - \widetilde{\mathbf{F}}$ verifying Eq. (3.5.13).
In fact, if $t \to \mathbf{x}(t)$, $t \in \mathcal{R}$, is a constraints compatible motion passing through $\boldsymbol{\xi}$ at $t = 0$ with speed $\boldsymbol{\eta}_2$, it will be $\widetilde{V}(\mathbf{x}(t)) = constant$; hence,

$$0 = -\frac{d\widetilde{V}(\mathbf{x}(t))}{dt}\Big|_{t=0} = -\frac{\partial \widetilde{V}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \cdot \dot{\mathbf{x}}(0) = \frac{\partial \widetilde{V}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \cdot \dot{\boldsymbol{\eta}}_2 = \widetilde{F}(\boldsymbol{\xi}) \cdot \boldsymbol{\eta}_2. \tag{3.5.16}$$

(3) The ambiguity seen in observation 2 has a physical interpretation: it is generally ambiguous to talk about the constraints reactions before having specified which are the other forces "not due to the constraints". Think of a point constrained to glide on a horizontal plane: we can always look at it as if it were subject to a force orthogonal to the plane and of arbitrary intensity G, besides the vertical downward gravity force $m\mathbf{g}$. The point will not change its motion, at least in absence of friction, but the reaction of the table will be $m\mathbf{g}$ upwards in the first case and $m\mathbf{g} + \mathbf{G}$ upwards in the second.
(4) Hence, on the basis of the above definition of ideality, the ideality of a system of constraints depends on the choice of $\mathbf{F}^{(a)}$: only once both $\mathbf{F}$ and $\mathbf{F}^{(a)}$ are given it is possible to define $\mathbf{R}$ and check Eq. (3.5.13). Therefore, the ideality of a constraint is not a property that can be described only in terms of the total force $\mathbf{F}$ producing it.
Translating into a mathematical model concrete problems, it often happens that one is given the constraints equations $\varphi^{(1)}, \ldots, \varphi^{(s)}$. and, separately, the active forces $\mathbf{F}^{(a)}$ and the "reaction of the constraints" $\mathbf{R}$. In fact, in applications it is often possible to distinguish operationally between the forces due to the constraints ("constraint reactions") and those due to other causes ("active forces"). In such cases, $\mathbf{R}$ is a priori given, or at least some of its properties are a priori given.
(5) Equation (3.5.13) is often called the "symbolic equation of dynamics" or "D'Alembert's principle" or "virtual works principle". The last name is usually given to Eq. (3.5.13) in its applications to statics where it is considered with $\boldsymbol{\eta}_1 = \mathbf{0}$ (see, also, the next comment and Observation 2, p.164, and the concluding remarks, p.241).
(6) Equation (3.5.13) is also read "the virtual work of an ideal constraint reaction always vanishes". This is perhaps the most suggestive way of reading

this equation since it stresses the fact that the velocity vector $\boldsymbol{\eta}_2$ is not the same as that, $\boldsymbol{\eta}_1$ provoking the reaction in $\boldsymbol{\xi}$. It is, in fact, the velocity of another possible motion through $\boldsymbol{\xi}$ ("virtual motion"). The word "work" is naturally a reference to the fact that $\mathbf{R}(\boldsymbol{\eta}_1, \boldsymbol{\xi}) \cdot \boldsymbol{\eta}_2$ is the work per unit time that the constraint reaction to the motion $\mathbf{x}$ passing at a given time through $\boldsymbol{\xi}$ with speed $\boldsymbol{\eta}_1$ performs on *another* motion passing, at the same time, through $\boldsymbol{\xi}$ with speed $\boldsymbol{\eta}_2$.

In the upcoming sections, we will analyze the physical meaning of Definition 8, i.e., we shall discuss the physical circumstances in which it becomes a relevant definition. Before that analysis, let us examine some remarkable consequences of this definition.

The first consequence is the following proposition: the "theorem of energy conservation for ideally constrained systems".

7 Proposition. *Let $t \to \mathbf{x}(t)$, $t \in I$, be a motion of a system of N point masses in $\mathcal{R}^d$ with masses $m_1, \ldots, m_N > 0$ subject to a system of s ideal holonomous constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$ and to a conservative active force $\mathbf{F}^{(a)}$ with potential energy $V^{(a)} \in C^\infty(\mathcal{R}^{Nd})$. Assume that $\mathbf{x}$ respects the constraints. Then there is a constant E such that*

$$T(t) + V^{(a)}(t) = E, \, t \in I, \tag{3.5.17}$$

where $V^{(a)}(t) = V^{(a)}(\mathbf{x}(t))$ and $T(t)$ is the kinetic energy of the motion $\mathbf{x}$ at time t.

Observation. The main point is that the above proposition does not assume that the reaction of the constraint is conservative in the sense of §3.1, but "only" that it is ideal, i.e., that it verifies Eq. (3.5.13). It can be velocity dependent, for instance.

Proof. It is an immediate consequence of the theorem of alive forces, proposition 2, §3.2, p.145, that the variation of kinetic energy between two times t_1, and t_2 is equal to the sum of the work performed on the motion $\mathbf{x}$ by the force $\mathbf{F}^{(a)}$, i.e., $V^{(a)}(t_1) - V^{(a)}(t_2)$, and by the reaction $\mathbf{R}$, given by

$$\sum_{j=1}^{N} \int_{t_1}^{t_2} \mathbf{R}^{(j)}(\dot{\mathbf{x}}(t), \mathbf{x}(t)) \, dt \tag{3.5.18}$$

However, by assumption, the motion $\mathbf{x}$ respects the constraints and, also, Eq. (3.5.13) holds. Using Eq. (3.5.13) with $\boldsymbol{\xi} = \mathbf{x}(t), \boldsymbol{\eta}_1 = \boldsymbol{\eta}_2 = \dot{\mathbf{x}}(t)$, we see that the work in Eq. (3.5.18) vanishes; hence, $T(t_2) - T(t_1) = V^{(a)}(t_1) - V^{(a)}(t_2)$, implying Eq. (3.5.17).

mbe

Far more interesting is the following proposition: the "least-action principle for ideally constrained systems".

8 Proposition. *Consider N points in $\mathcal{R}^d$, with masses $m_1, \ldots, m_N > 0$, subject to s holonomous ideal constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$ and to the active force $\mathbf{F}^{(a)}$, conservative, with potential energy $V^{(a)} \in C^\infty(\mathcal{R}^{Nd})$. Denote by $\mathbf{R}$ the constraint reaction. The action with Lagrangian*

$$\mathcal{L}(\boldsymbol{\eta}, \boldsymbol{\xi}, t) = \sum_{j=1}^{N} \frac{m_j}{2} \boldsymbol{\eta}^{(j)2} - V^{(a)}(\boldsymbol{\xi}) \tag{3.5.19}$$

is stationary in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$ on the motions which are generated by the force $\mathbf{F}^{(a)} + \mathbf{R} = \mathbf{F}$.
Furthermore, let $t \to x(t), t \in \mathcal{R}_+$, be a motion of the system developing under the action of the force $\mathbf{F}$ and respecting the constraints. Given $t_1 \in \mathcal{R}_+$, there exists $\overline{t} > t_1$ such that if $t_2 \in [t_1, \overline{t}]$, the action with Lagrangian (3.5.19) is locally minimal in $\mathcal{M}_{t_1,t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2) \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$ on the motion $\mathbf{x}$ observed for $t \in [t_1, t_2]$ and thought of as an element of $\mathcal{M}_{t_1,t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2) \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$.

Observations. (1) The importance of the above proposition lies in the fact that, if wisely used, it allows one to "eliminate" the degrees of freedom which are redundant because of the constraints. Suppose that one is able to find N C^∞ functions on $\mathcal{R}^\ell$ taking values in $\mathcal{R}^d$:

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_\ell) \to \mathbf{X}^{(i)}(\boldsymbol{\alpha}) = \mathbf{X}^{(i)}(\alpha_1, \ldots, \alpha_\ell), \tag{3.5.20}$$

$i = 1, \ldots, N$, such that, $\forall\, \boldsymbol{\alpha} \in \mathcal{R}^\ell$:

$$\varphi^{(j)}(\mathbf{X}(\boldsymbol{\alpha})) = 0, \qquad j = 1, \ldots, s \tag{3.5.21}$$

i.e., such that the image of $\mathcal{R}^\ell$ via the transformation (3.5.20) is a subset of $\mathcal{R}^{Nd}$ which automatically "verifies the constraints".[5]
Also, suppose one knows that the motion $\widehat{\mathbf{x}} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$ that we are studying and which develops under the action of the force $\mathbf{F}$, is the image in $\mathcal{R}^{Nd}$ of a motion $\mathbf{a} \in \mathcal{M}_{t_1,t_2}(\widehat{\mathbf{a}}(t_1), \widehat{\mathbf{a}}(t_2))$ in $\mathcal{R}^\ell$ via the transformation (3.5.20).
The above assumptions mean that we have a good understanding of the structure of the constraint so that we can find an explicit parametric representation of a class of configurations satisfying it.
Then the action $A(\mathbf{x})$ with Lagrangian (3.5.19) can be computed on the motions $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$ that are images via Eq.(3.5.21) of motions $\mathbf{a} \in \mathcal{M}_{t_1,t_2}(\widehat{\mathbf{a}}(t_1), \widehat{\mathbf{a}}(t_2))$ as $\widetilde{A}(\mathbf{a})$, where $\widetilde{A}$ is the action on $\mathcal{M}_{t_1,t_2}(\widehat{\mathbf{a}}(t_1), \widehat{\mathbf{a}}(t_2))$ with (t-independent) Lagrangian

[5] For instance, in the case of the point constrained on a line (§3.4), one can take $\ell = 1$ and $\alpha \to x(\alpha) = \xi(\alpha)$, $\alpha \in \mathcal{R}$, and the parameter α has the meaning of a curvilinear abscissa on the curve.

$$\widetilde{\mathcal{L}}(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \sum_{j=1}^{N} \frac{m_j}{2} \Big(\sum_{k=1}^{\ell} \frac{\partial \mathbf{X}^{(j)}(\boldsymbol{\alpha})}{\partial \alpha_k} \beta_k \Big)^2 - V^{(a)}(\mathbf{X}(\boldsymbol{\alpha})) \tag{3.5.22}$$

[see, also, §3.4, Eq. (3.4.5), where this is derived]. Since by Proposition 8 A is stationary in $\widehat{\mathbf{x}} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \,|\, \varphi^{(1)}, \ldots, \varphi^{(s)})$, then $\widetilde{A}$ is stationary in $\widehat{\mathbf{a}}$ on the entire set $\mathcal{M}_{t_1,t_2}(\widehat{\mathbf{a}}(t_1), \widehat{\mathbf{a}}(t_2))$ and, therefore, by Proposition 4, §3.3, this means that

$$\frac{d}{dt}\Big(\frac{\partial \widetilde{\mathcal{L}}}{\partial \beta_i}(\dot{\widetilde{\mathbf{a}}}(t), \widetilde{\mathbf{a}}(t))\Big) = \Big(\frac{\partial \widetilde{\mathcal{L}}}{\partial \alpha_i}(\dot{\widetilde{\mathbf{a}}}(t), \widetilde{\mathbf{a}}(t))\Big) \tag{3.5.23}$$

for $i = 1, 2, \ldots \ell$ and $t \in [t_1, t_2]$,

Equation (3.5.23) provide ℓ equations for the ℓ unknown functions $t \to \alpha_i(t)$, $i = 1, 2, ..., \ell$, $t \in [t_1, t_2]$. These are the equations of motion for the essential coordinates once the degrees of freedom which have become inessential because of the constraints have been eliminated.

It is of fundamental importance to realize the difference between the considerations of this section and those, apparently alike, of §3.4, pp.153-154. Those, in fact, had been developed assuming that the force $\mathbf{F}$ was conservative in the sense of §3.1. In the present case, as the example of §3.4 p.156 shows, the force will generally be velocity dependent. After a few exercises the reader will understand how great a simplification Eq. (3.5.23) implies in the deduction of the equations of motion, if compared to the alternative procedure of writing the equations of motions in the ordinary Cartesian coordinates followed by the elimination of the constraint reactions [remarkably absent in Eq. (3.5.23)] and of the redundant coordinates. In many instances, for example think of a rigid body, N can be large but ℓ very small.

(2) It is convenient to say a few words to explain why the name "principle" is granted to the Proposition 8 as well as to several other propositions or definitions already met (D'Alembert's principle, virtual work principle, etc.). Such names have interesting historical origins: the reader should not believe that the discussion of the laws of mechanics and the treatment of all the mechanical problems by the application of the equation $\mathbf{f} = m\mathbf{a}$, together with the two other laws of dynamics, to the point masses into which a system can be ideally decomposed has always been obvious and natural since the work of Newton.

As already remarked, Newton himself did not arrive in a very clear way to such a conclusion. For instance, in his study of rigid motions he had recourse to arguments quite different from modern methods based on the cardinal equations (i.e., on Newton's laws).

Both before and after Newton, philosophers were accustomed to studying mechanical problems on the basis of special assumptions, "principles", which were deduced by them through more or less general considerations often a bit obscure.

Newton's principles can be thought of as belonging to the above class of principles, and, initially, they were used particularly in the theory of heavenly bodies

motions. Together with the three principles first formulated by Newton, there already existed, more or less clearly formulated at least in particular cases, the energy conservation principle for simple systems (Huygens), the principle of the linear momentum conservation (going back at least to Descartes), the virtual works principle (which was used in the solution of problems in statics by Del Monte, Galilei, Stevin, etc.), the inertia principle (Galilei), and to these principles many others can be added: they were invented, even in years following Newton's time, to treat complex mechanical problems.
With Euler's work, the synthesis of all the different principles began through the realization that they could all be unified and deduced from Newton's and, what is often not sufficiently clearly stated, equivalent to Newton's if suitably interpreted (probably beyond the intentions and meanings the inventors attributed to them) (see, also, the concluding remarks to Chapter 3, p.241).

Let us go back to the simple proof of Proposition 8.

PROOF. Let $\mathbf{x} \in \mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \,|\, \varphi^{(1)}, \dots, \varphi^{(s)}) \equiv \mathcal{M}$, be a motion developing under the influence of $\mathbf{F} = \mathbf{F}^{(a)} + \mathbf{R}$. The action of $\mathbf{x}$ with respect to the Lagrangian (3.5.19) is

$$A(\mathbf{x}) = \int_{t_1}^{t_2} \Big\{ \sum_{j=1}^{N} \frac{m_j}{2} (\dot{\mathbf{x}}^{(j)}(t))^2 - V^{(a)}(\mathbf{x}(t)) \Big\}\, dt. \tag{3.5.24}$$

Let $\mathbf{y} \in \mathcal{V}_{\mathbf{x}}(\mathcal{M})$; let us compute the derivative with respect to ε of the function $A(\mathbf{y}_\varepsilon)$ in $\varepsilon = 0$. If we set (see Definition 5, §3.3) $\mathbf{z}(t) = \frac{\partial \mathbf{y}_\varepsilon(t,0)}{\partial \varepsilon}$ $= (\mathbf{z}^{(1)}(t), \dots, \mathbf{z}^{(N)}(t)), t \in [t_1, t_2]$, we have $\mathbf{z}(t_1) = \mathbf{z}(t_2) = \mathbf{0}$ and

$$\frac{d}{d\varepsilon} A(\mathbf{y}_\varepsilon)\big|_{\varepsilon=0} = \int_{t_1}^{t_2} \sum_{j=1}^{N} \Big\{ m_j \dot{\mathbf{x}}^{(j)}(t) \cdot \dot{\mathbf{z}}^{(j)}(t) - \frac{\partial V^{(a)}}{\partial \boldsymbol{\xi}^{(j)}}(\mathbf{x}(t)) \cdot \mathbf{z}^{(j)}(t) \Big\}\, dt. \tag{3.5.25}$$

By integrating the terms containing $\dot{\mathbf{z}}^{(j)}(t)$ by parts and using $\mathbf{z}(t_1) = \mathbf{z}(t_2) = \mathbf{0}$, one deduces, as usual,

$$\frac{d}{d\varepsilon} A(\mathbf{y}_\varepsilon)\big|_{\varepsilon=0} = -\int_{t_1}^{t_2} \sum_{j=1}^{N} \Big\{ m_j \ddot{\mathbf{x}}^{(j)}(t) + \frac{\partial V^{(a)}}{\partial \boldsymbol{\xi}^{(j)}}(\mathbf{x}(t)) \Big\} \cdot \mathbf{z}^{(j)}(t)\, dt. \tag{3.5.26}$$

The equations of motion for $\mathbf{x}$ are, by assumption,

$$m_j \ddot{\mathbf{x}}^{(j)}(t) = -\frac{\partial V^{(a)}}{\partial \boldsymbol{\xi}^{(j)}}(\mathbf{x}(t)) + \mathbf{R}^{(j)}(\dot{\mathbf{x}}(t), \mathbf{x}(t)); \tag{3.5.27}$$

hence, we cannot conclude that the right-hand side of Eq. (3.5.26) vanishes, but only that it is equal to

$$\frac{d}{d\varepsilon}A(\mathbf{y}_\varepsilon)\big|_{\varepsilon=0} = -\int_{t_1}^{t_2}\sum_{j=1}^{N}\mathbf{R}^{(j)}(\dot{\mathbf{x}}(t),\mathbf{x}(t))\cdot\mathbf{z}^{(j)}(t)\,dt. \tag{3.5.28}$$

However, Eq. (3.5.13) will allow us to infer that, $\forall\,\tau\in(t_1,t_2)$:

$$\sum_{j=1}^{N}\mathbf{R}^{(j)}(\dot{\mathbf{x}}(\tau),\mathbf{x}(\tau))\cdot\mathbf{z}^{(j)}(\tau)\equiv 0 \tag{3.5.29}$$

if we show that $(\mathbf{x}(t),\mathbf{z}(t))$ are a position-velocity pair compatible with the constraints; i.e., if we show the existence of a motion defined for t near τ and constraint compatible, which at $t=\tau$ is in $\mathbf{x}(t)$ with velocity $\mathbf{z}(t)$.

Recalling the definition of $\mathbf{z}$, one sees that such a motion indeed exists. To build it, one simply defines $t\to\mathbf{y}(\tau,t-\tau)$ for $t-\tau\in(-1,1)$, i.e., for t close to τ. This function of t has, for $t=\tau$, velocity $\mathbf{z}(\tau)$ and, furthermore, verifies the constraints and has a value $\mathbf{x}(\tau)$, for $t=\tau$, since $\mathbf{y}\in\mathcal{V}_{\mathbf{x}}(\mathcal{M})$.

We shall not explicitly prove the local minimum property: its (long) proof is entirely analogous to the proof of Proposition 37, §2.24, and should not present particular difficulties to the reader. mbe

3.5.1 Problems

1. Give an example of a holonomous constraint for a system of N point masses in $\mathcal{R}^3$ for which the only constraint compatible velocity $\boldsymbol{\eta}$ is $\boldsymbol{\eta}=\mathbf{0}$. (*Hint:* Find a constraint φ such that $\varphi(\boldsymbol{\xi})=0$ determines an isolated point.)

2.* Let $\varphi^{(1)},\ldots,\varphi^{(s)}$ be s holonomous constraints for a system of N point masses in R^3. Given a constraint compatible $\boldsymbol{\xi}\in R^{3N}$, assume that the s vectors $\frac{\partial\varphi^{(j)}(\boldsymbol{\xi})}{\partial\boldsymbol{\xi}}\in\mathcal{R}^{3N}$, $j=1,\ldots,s$, are linearly independent. Show that $\boldsymbol{\eta}$ is a constraint compatible velocity if and only if Eq. (3.5.9) holds. (*Hint:* The necessity is obvious. Conversely, consider the conditions on a constraint compatible motion of the form $t\to\boldsymbol{\xi}\,+\,t\,\boldsymbol{\eta}+\sum_{j=1}^{s}\delta_j(t)\frac{\partial\varphi^{(j)}}{\partial\boldsymbol{\xi}}$ given by

$$\varphi^{(k)}(\boldsymbol{\xi})\;+\;t\,\boldsymbol{\eta}+\sum_{j=1}^{s}\delta_j(t)\frac{\partial\varphi^{(j)}}{\partial\boldsymbol{\xi}}=0,\quad k=1,\ldots,s$$

which are regarded as equations for δ_j parameterized by t and solved, for $t=0$, by $\delta_j=0$. We now regard the left-hand side as a function of $t,\delta_1,\ldots,\delta_s$, and call it $\Phi^{(k)}(t,\delta_1,\ldots,\delta_s)$ and we try to define $\delta_j(t)$, $j=1,\ldots,s$, for t near zero, applying the implicit functions theorem (Appendix G). The Jacobian matrix for $t=0$, $\delta_1=\ldots=\delta_s=0$, is

$$M_{kh}=\frac{\partial\Phi^{(k)}}{\partial\delta_h}(\mathbf{0})=\sum_{p=1}^{3N}\frac{\partial\varphi^{(k)}}{\partial\xi_p}(\boldsymbol{\xi})\frac{\partial\varphi^{(h)}}{\partial\xi_p}(\boldsymbol{\xi}),\qquad h.k=1,\ldots,s$$

which has rank s, by the supposed linear independence of the vectors $\frac{\partial\varphi^{(k)}(\boldsymbol{\xi})}{\partial\boldsymbol{\xi}}$. In fact, the linear independence means that, $\forall\,\mathbf{c}\in\mathcal{R}^s$ $\sum_{k=1}^{s}c_k\frac{\partial\varphi^{(k)}(\boldsymbol{\xi})}{\partial\boldsymbol{\xi}}\neq\mathbf{0}$ unless $\mathbf{c}=\mathbf{0}$; therefore, if $\mathbf{c}\neq\mathbf{0}$:

$$\sum_{h,k=1}^{s}c_hc_kM_{hk}=\sum_{p=1}^{3N}\sum_{h,k=1}^{s}c_hc_k\frac{\partial\varphi^{(k)}}{\partial\xi_p}(\boldsymbol{\xi})\frac{\partial\varphi^{(h)}}{\partial\xi_p}(\boldsymbol{\xi})=\sum_{p=1}^{3N}\Big(\sum_{h=1}^{s}c_h\frac{\partial\varphi^{(h)}}{\partial\xi_p}(\boldsymbol{\xi})\Big)^2>0$$

and, since $M_{hk} = M_{kh}$, this means that the matrix M is positive definite; hence $\det M > 0$ (see Appendix F). Thus, by the implicit functions theorem, there exist s C^∞ functions $t \to \delta_j(t)$, $j = 1, \ldots, s$, defined near $t = 0$ such that the motion $t \to \boldsymbol{\xi} + t\,\boldsymbol{\eta} + \sum_{j=1}^{s} \delta_j(t) \frac{\partial \varphi^{(j)}}{\partial \boldsymbol{\xi}}$ verifies the constraints and, furthermore, $\delta_j(0) = 0$ and, by the implicit functions theorem:

$$\delta_k(t) = -t \sum_{s=1}^{s} (M^{-1})_{kh} \frac{\partial \varphi^{(h)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \cdot \boldsymbol{\eta} + t^2\, \widetilde{\delta}_k(t)$$

for some C^∞ functions $\widetilde{\delta}_k(t)$ defined near $t = 0$. By the assumption on $\boldsymbol{\eta}$, the t-linear term vanishes: hence, $\dot{\delta}_k(0) = 0$, i.e., $\dot{\mathbf{x}}(0) = \boldsymbol{\eta}$ ... .).

3.* Given a system of N points in $\mathcal{R}^3$, with masses $m_1, \ldots, m_N > 0$, subject to ideal holonomous constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$ and to active force $\mathbf{F}^{(a)}$, show the possibility of an explicit expression for the reaction $\mathcal{R}$ acting on a constraint compatible motion, at a time t when $\mathbf{x}(t) = \boldsymbol{\xi}, \dot{\mathbf{x}}(t) = \boldsymbol{\eta}$. (*Hint:* Let $m_1 = \ldots = m_N = 1$ for simplicity and suppose also that the s vectors in $\mathcal{R}^{3N}$, $\frac{\partial \varphi^{(j)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}$, $j = 1, \ldots, s$, are independent, again for simplicity. From $\varphi^{(j)}(\mathbf{x}(t)) \equiv 0$, $j = 1, \ldots, s$, deduce by two-fold differentiation

$$\frac{\partial \varphi^{(j)}}{\partial \boldsymbol{\xi}}(\mathbf{x}(t)) \cdot \ddot{\mathbf{x}}(t) + \sum_{p,q=1}^{N} \frac{\partial^2 \varphi^{(j)}}{\partial \xi_p \partial \xi_q}(\mathbf{x}(t)) \dot{x}_p(t) \dot{x}_q(t) \equiv 0$$

and then combine this equation with the equation of motion $\ddot{\mathbf{x}} = \mathbf{F}^{(a)} + \mathbf{R}$ to obtain

$$\frac{\partial \varphi^{(j)}}{\partial \boldsymbol{\xi}}(\boldsymbol{\xi}) \cdot \mathbf{R}(\boldsymbol{\eta}, \boldsymbol{\xi}) = -\left\{ \frac{\partial \varphi^{(j)}}{\partial \boldsymbol{\xi}}(\boldsymbol{\xi}) \cdot \mathbf{F}^{(a)}(\boldsymbol{\xi}) + \sum_{p,q=1}^{N} \frac{\partial^2 \varphi^{(j)}}{\partial \xi_p \partial \xi_q}(\boldsymbol{\xi}) \eta_p \eta_q \right\},$$

$j = 1, \ldots, s$. But, by the ideality assumption, $\mathbf{R}(\boldsymbol{\eta}, \boldsymbol{\xi})$ has to be orthogonal to every $\boldsymbol{\eta}'$ such that $\frac{\partial \varphi^{(j)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \cdot \boldsymbol{\eta}' = 0$ [see Problem 2 and Eq. (3.5.13)]. Hence, $\mathbf{R}$ has to be a linear combination of the s vectors $\frac{\partial \varphi^{(j)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}$, $j = 1, \ldots, s$, and the coefficients can be determined by the scalar products $\frac{\partial \varphi^{(j)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \cdot \mathbf{R}$, $j = 1, \ldots, s$, since the s vectors $\frac{\partial \varphi^{(j)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}$ are linearly independent. Deal also with the general case: different masses and linearly dependent vectors $\frac{\partial \varphi^{(j)}(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}$.)

4. A "constraint" of the form $\varphi(\boldsymbol{\xi}) \geq 0$ for a system of N point masses in $\mathcal{R}^d$ is called "unilateral". Show that such constraints are not more general than those considered in Definition 7. (Hint: Let $\alpha \to \chi(\alpha)$ be a C^∞ function, strictly positive if $\alpha < 0$ and zero if $\alpha \geq 0$; then consider the constraint $\psi(\boldsymbol{\xi}) = \chi(\varphi(\boldsymbol{\xi})) = 0$, etc.)

5. Show that any velocity is compatible with a unilateral constraint $\varphi \geq 0$ in the positions $\boldsymbol{\xi}$, where $\varphi(\boldsymbol{\xi}) > 0$.

6. Which are the velocities $\boldsymbol{\eta}$ compatible with a unilateral constraint $\varphi(\boldsymbol{\xi}) \geq 0$ in a position $\boldsymbol{\xi}$ where $\varphi(\boldsymbol{\xi}) = 0$? Suppose $\frac{\partial \varphi(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \neq \mathbf{0}$. (*Answer:* Those such that $\boldsymbol{\eta} \cdot \frac{\partial \varphi(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = 0$, i.e., the same as those for the constraint $\varphi(\boldsymbol{\xi}) = 0$!)

7. Extend the notion of velocity $\boldsymbol{\eta}$ compatible, in a configuration $\boldsymbol{\xi}$, with some holonomous (unilateral or not) constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$ by saying that $\boldsymbol{\eta}$ is constraint compatible at $\boldsymbol{\xi}$ if there is a constraint compatible motion $t \to \mathbf{x}(t)$, defined for $t > 0$ small enough (rather than for $|t|$ small enough) such that $\dot{\mathbf{x}}(0) = \boldsymbol{\eta}, \mathbf{x}(0) = \boldsymbol{\xi}$. Show that there are cases where $\boldsymbol{\eta}$ can be constraint compatible in this new sense without being so in the old one of Definition 7. We call the velocities which are constraint compatible in the new sense "(+)-compatible velocities". (*Hint:* The two notions will differ when $\boldsymbol{\xi}$ is such that $\varphi(\boldsymbol{\xi}) = 0$ in the case of a

unilateral constraint $\varphi \geq 0$. Give a physical interpretation of such extra velocities in terms of "collision velocities" with the constraint.)

8.* Show that the smoothness requirements on $\mathbf{F}$ and $\mathbf{F}^{(a)}$ used for the Definition 8 of an ideal constraint cannot generally hold if the system is subject to a unilateral constraint $\varphi \geq 0$ (thought of as a constraint via the construction of Problem 4); i.e., a unilateral constraint cannot, in general, be ideal in the sense of Definition 8. (*Hint:* There are, in general, motions starting in the region $\varphi > 0$ which in a finite time reach a point $\boldsymbol{\xi}_0$, where $\varphi(\boldsymbol{\xi}_0) = 0$, "collision with the constraint", with a speed which is not (+)-constraint compatible in the sense of Problem 7. At this point the speed must have a discontinuity against the assumption that $\mathbf{F} \in C^\infty(\mathcal{R}^{3N})$ and the regularity theorem for the differential equations.)

3.6 Real and Ideal Constraints

The discussion of §3.5 is largely unsatisfactory.

The notion of constraints used there has been given on a purely mathematical basis and it is quite unclear which is the physical phenomenon mathematically modeled by the constraints, ideal or not, of the preceding sections.

In this and in the following sections, we will radically modify the point of view to show that an ideal constraint for a system of N point masses can also be thought of as a limiting case of suitable very strong conservative force fields which oblige the trajectories to lie on certain surfaces in $\mathcal{R}^{Nd}$ or in their vicinity.

From a physical viewpoint, one always imagines a constraint as a complex of forces acting on a system of point masses and due to their tendency to deform some obstacles. Such a tendency provokes imperceptibly small (at least as far as our observations are concerned[6]) deformations of the obstacles. Think of a point constrained on a rail or on a surface, or think of a rigid system.

Note, also, that in the above concrete cases, the elegant theory of §3.5 is totally useless: the constraints now constrain in an approximate sense only and, therefore, they are not of the type considered there.

The question which is more interesting for us in this context is whether or not the solutions of the equations obtained by minimizing the Lagrangian (3.5.19) on the motions constrained by s holonomous constraints $\varphi^{(1)}, \ldots, \varphi^{(s)}$ (see Definition 8, §3.5) provide good approximations to the real motion under the influence of the real constraints, which necessarily constrain only in an approximate sense.

This is a really interesting problem in physics and applications, in contrast with the question underlying §3.5 which, abstractly, asked for a definition of a perfect constraint that would give rise to a sufficient condition in order that the equations of motion could be deduced from the least-action principle

[6] When they can be appreciated, one no longer speaks of a constraint.

associated with the Lagrangian (3.5.19), for the $(\varphi^{(1)},\ldots,\varphi^{(n)})$-constrained motions (see §3.5, Proposition 8, p.163).

To understand better the spirit and the meaning of the various definitions that will follow, it is convenient to analyze a simple but significant example.

Consider a point with mass $m>0$ in $\mathcal{R}^2$ and suppose that it is subject to an elastic force with potential energy $V^{(a)}=\frac{m\omega^2}{2}(x^2+y^2)$ and to a restoring conservative force toward the $y=0$ axis with potential energy

$$\lambda\, W(x,y)=\frac{\lambda\, m}{2}\,y^2 \tag{3.6.1}$$

Consider the motions under the action of the force with potential energy

$$V^{(a)}+\lambda\, W\;=\;\frac{m\omega^2}{2}(x^2+y^2)+\frac{\lambda\, m}{2}\,y^2 \tag{3.6.2}$$

It is intuitively clear that if A is very large, such a force simulates a constraint to the line $y=0$ in a sense which has still to be precisely understood. For this purpose study the motions which start on the $y=0$ axis and develop under the influence of the force with potential energy of Eq. (3.6.2). The equations of motion are

$$\begin{aligned}&m\ddot{x}=-\omega^2 m\,x, \qquad m\ddot{y}=-\omega^2 m\,y-m\lambda y,\\ &x(0)=x_0.\ \dot{x}(0)=v_0,\ \ y(0)=0,\ \ \dot{y}_0=w_0,\end{aligned} \tag{3.6.3}$$

and, to be definite, suppose $x_0>0$.

Because of their extreme simplicity, Eqs. (3.6.3) can be elementarily solved: if $t\to(x(t),y(t))$, $t\in\mathcal{R}$, denotes the solution of Eqs. (3.6.3):

$$x_\lambda(t)=\sqrt{x_0^2+\frac{v_0^2}{\omega^2}}\cos(\omega t+\varphi_0), \qquad y_\lambda(t)=\frac{w_0}{\sqrt{\lambda+\omega^2}}\sin(\omega t+\varphi_0), \tag{3.6.4}$$

with $\varphi_0=\operatorname{arctg}-\frac{v_0}{\omega x_0}$. One then sees that the limit as $\lambda\to\infty$ of the motion of Eqs. (3.6.4) is the motion $t\to(x(t),y(t))$ with

$$x(t)=\big(x_0^2+\frac{v_0^2}{\omega^2}\big)^{\frac12}\cos(\omega t+\varphi_0), \quad y(t)\equiv 0 \tag{3.6.5}$$

for all w_0. This is exactly the solution of the equations obtained by imposing stationarity on the motions constrained by the ideal holonomous constraint $\xi_2=0$ for the action with Lagrangian:

$$\mathcal{L}(\eta_1,\eta_2,\xi_1,\xi_2)=\frac{m}{2}(\eta_1^2+\eta_2^2)-\frac{m\omega^2}{2}(\xi_1^2+\xi_2^2) \tag{3.6.6}$$

On the basis of Observation (1) to Proposition 8 of §3.5, these equations coincide with those for the motions $t\to x(t)$, $t\in\mathcal{R}$, in $\mathcal{R}^1$ which make stationary the "constrained Lagrangian":

$$\widetilde{\mathcal{L}} = \frac{m}{2}\eta_1^2 - \frac{m\omega^2}{2}\xi_1^2 \tag{3.6.7}$$

which, in our case, is what Eq. (3.6.6) becomes by imposing the constraint $\xi_2 = 0$.

It is interesting to note that the more "rigid" the approximate constraint realized by Eq. (3.6.2), the smaller the deviations from a motion respecting the constraints ($|y_\lambda| \le O(\frac{1}{\sqrt{\lambda}})$) for λ large [see Eqs. (3.6.4)]. At the same time, however, the coordinate $y_\lambda(t)$, which simply represents the violation of the constraint, oscillates more and more rapidly: in fact, its vibrations have a frequency:

$$\nu_\lambda = \frac{\sqrt{\lambda + \omega^2}}{2\pi} \tag{3.6.8}$$

These very small but very-high-frequency vibrations ("fatigue vibrations" of the constraint) provide a good intuitive representation of the effect of an approximate ideal constraint on the motion of a system. In general, it is possible to think that a system of N point masses subject to an approximate ideal constraint moves as if it were on the surface $\Sigma \subset \mathcal{R}^{Nd} \subset \mathcal{R}^{Nd}$ defined by the constraint, with some very small elongations orthogonal to Σ: described by oscillatory motions with very small amplitude and very large frequency.

On the basis of the above heuristic discussion, the following definition should appear quite natural.

9 Definition. *Given a system of N point masses, with masses $m_1, \ldots, m_N$ in $\mathcal{R}^d$, let $\Sigma \subset \mathcal{R}^{Nd}$ be a closed set and let W be a real $C^\infty(\mathcal{R}^{Nd})$ function vanishing on Σ and having there a strict minimum; i.e., Σ is the set of the points $\boldsymbol{\xi} \in \mathcal{R}^{Nd}$ where $W(\boldsymbol{\xi}) = 0$ and for all $\boldsymbol{\xi} \notin \Sigma$ it is $W(\boldsymbol{\xi}) > 0$. We shall say that the conservative force law with potential energy*

$$\boldsymbol{\xi} \to \lambda\, W(\boldsymbol{\xi}) \ge 0. \qquad \lambda > 0 \tag{3.6.9}$$

is a "model of conservative approximate constraint to the region Σ with structure W and rigidity λ".
We shall denote such a force law by (Σ, W, λ). If Σ is a regular surface with dimension $\ell < Nd$, we shall say that the constraint model (Σ, W, λ) is a bilateral approximate conservative constraint "with dimension ℓ" or "with codimension $Nd - \ell$", (see also Definition 10 below).

Observation. In general Σ may contain interior points: in this case, one says that Eq. (3.6.9) is a model for a "unilateral" approximate constraint.

It is convenient to recall the definition of a regular surface in $\mathcal{R}^d$.

10 Definition. *Let $\boldsymbol{\beta} \to \boldsymbol{\Xi}(\boldsymbol{\beta})$ be a C^∞ function defined on a convex open $\Omega \subset \mathcal{R}^d$, taking its values in a neighborhood $U \subset \mathcal{R}^d$. Suppose that $\boldsymbol{\Xi}$ is invertible, i.e., one to one, as a map of Ω onto U and, furthermore, assume that $\boldsymbol{\xi}$ is nonsingular, i.e., that its Jacobian matrix defined by*

$$J(\boldsymbol{\beta}) = \frac{\partial \Xi_i}{\partial \beta_j}(\boldsymbol{\beta}), \qquad j = 1, \ldots, d \tag{3.6.10}$$

has a non vanishing determinant, $\forall \boldsymbol{\beta} \in \Omega$.
We shall say that $\boldsymbol{\Xi}$ *"establishes a regular system of coordinates on* U*" and, if* $\boldsymbol{\xi} = \boldsymbol{\Xi}(\boldsymbol{\beta})$*, we shall say that* $\boldsymbol{\beta} \in \Omega$ *is the coordinate of* $\boldsymbol{\xi}$ *in the coordinate system on* U *associated with* $\boldsymbol{\Xi}$*. The just-described coordinate system will be denoted* $(U, \boldsymbol{\Xi})$ *and* Ω *will be called the "basis" of the coordinate system.*
A closed set $\Sigma \subset \mathcal{R}^d$ *will be called a "regular* ℓ*-dimensional surface" if for all* $\boldsymbol{\xi}_0 \in \Sigma$ *it is possible to find a neighborhood* U_0 *and a regular system of coordinates* $(U_0, \boldsymbol{\Xi})$ *with basis* Ω_0 *such that the points of* $\Sigma \cap U_0$ *are all those with coordinates* $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_\ell)$ *such that*

$$\beta_1 = \beta_2 = \ldots, \beta_\ell = 0 \tag{3.6.11}$$

We say that $(U_0, \boldsymbol{\Xi})$ *is a local system of coordinates "adapted to* Σ*".*
A regular s*-dimensional surface is also called a regular surface of codimension* $d - s$*.*

Going back to the definition of approximate conservative constraint attention will be confined, from now on, to approximate bilateral conservative constraints with dimension ℓ, or, as it is customary to say, with "ℓ degrees of freedom", $0 < \ell < Nd$.

Consider a system of N points in R^d, with masses $m_1, \ldots, m_N > 0$, subject to the action of a conservative force with potential energy $V^{(a)} \in C^\infty(\mathcal{R}^{Nd})$, bounded from below, and to a model of approximate conservative constraint (Σ, W, λ), bilateral and ℓ-dimensional.

Suppose that Σ is defined by equations of the type

$$\varphi^{(1)}(\boldsymbol{\xi}) = \ldots = \varphi^{(s')}(\boldsymbol{\xi}) = 0 \tag{3.6.12}$$

with $\varphi^{(i)} \in C^\infty(\mathcal{R}^{Nd})$ (the number s' need not be $(Nd - \ell)$, although this will often be the case). It is natural to study the motions $t \to \mathbf{x}(t)$, $t \in \mathcal{R}_+$, developing under the action of the conservative force with potential energy

$$\boldsymbol{\xi} \to V^{(a)}(\boldsymbol{\xi}) + \lambda W(\boldsymbol{\xi}), \tag{3.6.13}$$

following an initial datum $\mathbf{x}_\lambda(0) = \boldsymbol{\xi}_0$, $\dot{\mathbf{x}}_\lambda(0) = \boldsymbol{\eta}_0$ with $\boldsymbol{\xi}_0$ compatible with the constraints

$$\boldsymbol{\xi}_0 \in \Sigma. \tag{3.6.14}$$

The question is whether there exists a limit

$$\mathbf{x}(t) = \lim_{\lambda \to +\infty} \mathbf{x}_\lambda(t), \qquad t \in \mathcal{R}_+. \tag{3.6.15}$$

Furthermore, one asks if $t \to \mathbf{x}(t)$, $t \in \mathcal{R}_+$, coincides (when existing) with a motion developing under the action of the ideal constraints of Eq. (3.6.12)

and of the active force with potential $V^{(a)}$, in conformity with the Definition 8, p.160, and Proposition 8, p.163.

It is easy to realize that there cannot be a positive answer if the problem is posed in the above generality.

Just reconsider the point mass in $\mathcal{R}^2$, with mass $m > 0$, constrained to the line $\xi_2 = 0$ with the new constraint model (Σ, W, λ) with $\Sigma = \{\xi_1\}$-axis and, if $\xi_1 \equiv x, \xi_2 \equiv y$,

$$W(x, y) = \frac{m}{2}\, y^2\, (1 + x^2) \tag{3.6.16}$$

subject, also, to the same active force with potential energy $V^{(a)} = \frac{m\omega^2}{2}(x^2 + y^2)$. The equations of motion, similar to Eqs. (3.6.3), now become

$$m\, \ddot{x} = -m\, \omega^2 x - \lambda\, m\, y^2 x, \qquad m\, \ddot{y} = -\lambda\, m\, y\, (1 + x^2) - m\, \omega^2\, y, \tag{3.6.17}$$

with $x(0) = x_0$, $\dot{x}(0) = v_0$, $y(0) = 0$, $\dot{y}(0) = w_0$. These equations are more complex than Eqs. (3.6.3), and will be discusses only in a heuristic, non rigorous way.

Let $t \to (x(t), y(t))$, $t \in \mathcal{R}_+$, be the solution of Eqs. (3.6.17). Energy conservation implies that if

$$E = \frac{m}{2}(v_0^2 + w_0^2) + \frac{m\omega^2}{2}\, x_0^2, \tag{3.6.18}$$

one has, $\forall\, t > 0$,

$$E = m\frac{\dot{x}_\lambda(t)^2 + \dot{y}_\lambda(t)^2}{2} + m\omega^2\frac{x_\lambda(t)^2 + y_\lambda(t)^2}{2} + \lambda m\frac{y_\lambda(t)^2(1 + x_\lambda(t)^2)}{2} \tag{3.6.19}$$

Then Eq. (3.6.19) implies

$$|\dot{x}_\lambda(t)|, |\dot{y}_\lambda(t)| \le \left(\frac{2E}{m}\right)^{\frac{1}{2}}, \tag{3.6.20}$$

$$|x_\lambda(t)| \le \left(\frac{2E}{m\omega^2}\right)^{\frac{1}{2}}, \tag{3.6.21}$$

$$|y_\lambda(t)| \le \left(\frac{2E}{m\lambda}\right)^{\frac{1}{2}} \tag{3.6.22}$$

which follow by observing that all the addends in Eq. (3.6.19) are nonnegative.

Fix a finite time interval $[0, T]$ and note that the first of Eqs. (3.6.17) together with Eqs. (3.6.20), (3.6.21), and (3.6.22) implies that the function $\ddot{x}_\lambda(t)$ has a uniformly bounded modulus for all $\lambda > 1$ and for $t \in [0, T]$. Hence, if T is "small", one can think that the function $t \to x_\lambda(t)$, $t \in [0, T]$, is practically constant together with its first derivative and, then, heuristically set $x_\lambda(t) = x_0$, $t \in [0, T]$ in the second equation of Eqs. (3.6.17). Within this "approximation", Eqs. (3.6.17) becomes "elementarily soluble":

$$y_\lambda(t) \simeq \frac{w_0}{(\omega^2 + \lambda\,(1+x_0^2))^{\frac{1}{2}}} \sin(\omega^2 + \lambda\,(1+x_0^2)^{\frac{1}{2}})t \tag{3.6.23}$$

which shows that, at least if $t \in [0,T]$,, $y_\lambda(t)$ varies very quickly if λ is large, with a frequency $\nu_\lambda \simeq (\omega^2 + \lambda\,(1+x_0^2))^{\frac{1}{2}}$ remaining, nevertheless, small by Eq. (3.6.22).

Since, as just seen, $x_\lambda(t)$ varies slowly (essentially λ independently), we can substitute in the first of Eqs. (3.6.17) $\lambda\, y_\lambda(t)^2$ with its average value $\overline{\lambda\, y_\lambda(t)^2}$ between 0 and t, if t is large compared to the characteristic time T_λ of variation of y, namely $T_\lambda \simeq 2\pi O(\lambda^{-\frac{1}{2}})$:

$$\begin{aligned}\overline{\lambda\, y_\lambda(t)^2} \simeq& \frac{\lambda}{t}\int_0^t \frac{w_0^2\,\big(\sin(\omega^2+\lambda\,(1+x_0^2)^{\frac{1}{2}})\tau\big)^2}{\omega^2+\lambda\,(1+x_0^2)}d\tau \\ =& \frac{\lambda\, w_0^2}{\omega^2+\lambda\,(1+x_0^2)}\Big[\frac{\int_0^{t\sqrt{\omega^2+\lambda\,(1+x_0^2)}}(\sin\theta)^2\, d\theta}{t\,\sqrt{\omega^2+\lambda\,(1+x_0^2)}}\Big]\end{aligned} \tag{3.6.24}$$

where Eq. (3.6.23) has been used and θ has been defined as $\theta = \tau\,(\omega^2+\lambda\,(1+x_0^2))^{\frac{1}{2}} = \frac{2\pi\tau}{T_\lambda}$. By assumption $t \gg T_\lambda = 2\pi(\omega^2+\lambda\,(1+x_0^2)^{-\frac{1}{2}})$, therefore the integral in square brackets can be replaced by

$$\lim_{R\to\infty}\frac{1}{R}\int_0^R (\sin\theta)^2 d\theta = \frac{1}{2}. \qquad \text{Hence,} \tag{3.6.25}$$

$$\overline{\lambda\, y_\lambda(t)^2} \simeq \frac{1}{2}\frac{\lambda\, w_0^2}{\omega^2+\lambda\,(1+x_0^2)} \simeq \frac{1}{2}\frac{w_0^2}{1+x_0^2} \tag{3.6.26}$$

if λ is large enough. Then substituting $\lambda\, y_\lambda(t)^2 \to \overline{\lambda\, y_\lambda(t)^2}$ in the first of Eqs. (3.6.17), one finds

$$m\ddot{x}_\lambda = -m\omega^2\, x_\lambda - \frac{w_0^2}{2(1+x_0^2)}\, x_\lambda \tag{3.6.27}$$

for t near 0 (but $t \gg T_\lambda$) and λ large (note that $T_\lambda \to 0$ as $\lambda \to +\infty$).

For arbitrary values of t, a similar argument suggests that, in general, the acceleration $\ddot{x}_\lambda$ should verify the equation (when $\lambda \to +\infty$):

$$\ddot{x} = -m\omega^2\, x - \frac{mw_0^2}{2}\frac{x}{\sqrt{1+x^2}\sqrt{1+x_0^2}}. \tag{3.6.28}$$

Hence, the model of constraint to the line $y = 0$ with the structure of Eq. (3.6.16) does not give rise to the motions that develop under the action of an ideal constraint to the line $y = 0$ and of an active force with potential energy $V^{(a)}(x) = \frac{1}{2}m\omega^2 x^2$, when $\lambda \to +\infty$, as one could have naively expected.

Rather one should think that the limit motion, for $\lambda \to +\infty$, of x_λ is a motion subject to an ideal constraint to $y = 0$ and to the active force whose potential energy is

$$V'^{(a)}(x) = \frac{1}{2}m\omega^2 x^2 + \frac{w_0^2 m}{4} 2\sqrt{\frac{1+x^2}{1+x_0^2}} \tag{3.6.29}$$

which depends on the initial velocity w_0 transversal to the constraint (and, of course, on the particular form of the structure function W).

It is then possible to think, in general, that in the limit of infinite rigidity, the model of a conservative bilateral approximate constraint generates motions which respect the constraints and develop as if they were ideal, but under the influence of an active force modified with respect to the one with potential energy $V^{(a)}$ which naively could be thought to be the force "not due to the constraints". In general, the structure W of the constraints has some influence, even for λ large, and contributes to the active forces in a way that may also depend on the initial data or, better, on the "initial stresses on the constraints", as in the case of the last example, where the active force depends also on the initial velocity component w_0 orthogonal to the constraint. This conjecture also sheds some light on the slightly formal distinction in §3.5 between the active force and the reaction of the constraint.

In the following section we will deal with questions related to the following problems.

(1) Which further condition is it necessary to place on an approximate constraint model (Σ, W, λ) to imply that the motion $t \to \mathbf{x}_\lambda(t)$, $t \geq 0$, developing under the action of the force with potential energy

$$V^{(a)} + \lambda W, \tag{3.6.30}$$

and following the initial datum

$$\mathbf{x}(0) = \boldsymbol{\xi}_0 \in \Sigma, \qquad \dot{\mathbf{x}}(0) = \boldsymbol{\eta}_0 \tag{3.6.31}$$

is well approximated by the motion that takes place under the action of the active force with potential energy $V^{(a)}$ and of the ideal constraints $\varphi^{(1)}, \ldots, \varphi^{(s')}$ and follows the initial datum $\mathbf{x}(0) = \boldsymbol{\xi}_0$, $\dot{\mathbf{x}}(0) = \boldsymbol{\eta}_0^\Sigma$, where $\boldsymbol{\eta}_0^\Sigma$ is a suitable "projection of $\boldsymbol{\eta}_0$ on Σ", assuming that Σ is determined by the equations $\varphi^{(1)}(\boldsymbol{\xi}) = \ldots = \varphi^{(s')}(\boldsymbol{\xi}) = 0$? (See Definitions 7 and 8, §3.5.)
In other words the question is: when does an approximate conservative constraint appear as well approximated by an ideal constraint model in the sense of Definition 8, §3.5, with the "naive" identification of the active forces?

(2) If (Σ, W, λ) is a model of an approximate conservative constraint, is it true that the motion developing under the action of a force with the potential energy of Eq. (3.6.30) and following the initial datum of Eq. (3.6.31) is well approximated, as $\lambda \to +\infty$, by a motion developing under the influence of the ideal constraints $\varphi^{(1)}, \ldots, \varphi^{(s')}$ (determining Σ) and of a conservative active force with potential energy $V'^{(a)}$, possibly different from $V^{(a)}$ and $(\boldsymbol{\eta}_0, \boldsymbol{\xi}_0)$-dependent?

(3) In the same situation as that of question (2) and if $\boldsymbol{\eta}_0$ is suitable, i.e. $\boldsymbol{\eta}_0 \equiv \boldsymbol{\eta}_0^{\Sigma}$, where $\boldsymbol{\eta}_0^{\Sigma}$ is as in question (1), is it true that $V'^{(a)} = V^{(a)}$? [This seems to be true in the example, heuristically explicitly studied above, about the constraint to the line $y = 0$ generated by Eq. (3.6.16), when $w_0 = 0$; see Eqs. (3.6.17) and (3.6.29)].

Actually, we shall really study in detail question (1) only, which we shall refer to as the problem of the determination of "sufficient perfection conditions for approximate bilateral conservative constraints".

It will be useful and necessary to analyze in some deeper way the kinematics of the system of point masses subject to constraints: this is a purely geometric analysis, very suggestive for its relationship with differential geometry. The following section is mainly devoted to this task.

3.6.1 Exercises and Problems

1. Show that the polar coordinates are a regular system of coordinates in various regions $U \subset \mathcal{R}^2$ or $U \subset \mathcal{R}^3$.

2. Show that the surface of a sphere is a regular surface in the sense of Definition 10, p.170.

3. Show that the surface of the paraboloid $z = (x^2 + y^2)/2$ is a regular two-dimensional surface in the sense of Definition 10, p.170. Treat similarly the hyperboloid and ellipsoid cases.

4. Consider the ellipsoid surface $\frac{x^2}{a} + \frac{y^2}{b} + \frac{z^2}{c} = 1$, $0 < a < b < c$, and show that if $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3) \equiv (w, u, v)$

$$
\begin{aligned}
x =& (1+w)\sqrt{a\,\frac{(u-a)(v-a)}{(b-a)(c-a)}}\\
y =& (1+w)\sqrt{b\,\frac{(u-b)(v-b)}{(a-b)(c-b)}}\\
x =& (1+w)\sqrt{a\,\frac{(u-c)(v-c)}{(a-c)(b-c)}}
\end{aligned}
$$

is a local system of regular coordinates in the vicinity of various points of the ellipsoid's surface. This system is adapted to the surface itself, which has equations $\beta_1 \equiv w = 0$ ("Jacobi's coordinates"). The domains $w = 0, u \in (a, b), v \in (b, c)$ and $w = 0, u \in (b, c), v \in (a, b)$ give the part of the ellipsoid situated in the first octant in $\mathcal{R}^3$ with the exception of a few lines; determine the latter lines.

5. Let $V \in C^\infty(\mathcal{R})$, $V'(\xi) \equiv \frac{dV(\xi)}{d\xi} \neq 0$, $\forall \xi \neq 0$, $\lim_{\xi \to \pm\infty} V'(\xi) = +\infty$, $V(0) = 0$. Given a point $(\eta, \xi) \in \mathcal{R}^2$, $(\eta, \xi) \neq (0, 0)$, define

$$
E = \frac{\eta^2}{2} + V(\xi), \qquad T(E) = 2\int_{x_-(E)}^{x_+(E)} \frac{d\xi'}{\sqrt{2(E - V(\xi'))}},
$$

where $x_-(E) < x_+(E)$ are the two roots of $E - V(\xi) = 0$. Define

$$\varphi(\eta,\xi)=\frac{2\pi}{T(E)}\cdot\left\{\begin{array}{l}\text{time necessary to the motion } x \text{ such that}\\ \ddot{x}=-dV/d\xi \text{ with initial datum } (0,x_(E))\\ \text{to reach } (\eta,\xi)\end{array}\right\}$$

which defines φ (mod 2π) (as every motion with initial velocity η and position ξ is periodic and, sooner or later, visits $x_-(E)$ with velocity 0).
Show that the coordinates (E,φ) are a regular system of coordinates near any point in $\mathcal{R}^2$ other than the origin, and that they generalize the polar coordinates (ϱ,θ) with E being analogous to $\varrho^2/2$ and φ to θ. (*Hint:* First consider the case $V(\xi)=\xi^2/2$ and explicitly find $\varphi(\eta,\xi)$. Draw the qualitative form of the curves $\eta^2/2+V(\xi)=E$ as E varies.)

3.7 Kinematics of Quasi-constrained Systems. Reformulation of Perfection Criteria for Approximate Conservative Constraints

In this section $\mathcal{R}^{Nd}$ is regarded as a vector space in which the scalar product between two vectors $\boldsymbol{\eta}=(\boldsymbol{\eta}^{(1)},\ldots,\boldsymbol{\eta}^{(N)})$ and $\boldsymbol{\chi}=(\boldsymbol{\chi}^{(1)},\ldots,\boldsymbol{\chi}^{(N)})$ is defined by

$$\sum_{i=1}^{N} m_i\boldsymbol{\eta}^{(i)}\cdot\boldsymbol{\chi}^{(i)}, \tag{3.7.1}$$

where $m_1,\ldots,m_N$ are given positive numbers. The length of a vector is then

$$||\boldsymbol{\eta}||=\Big(\sum_{i=1}^{N} m_i{\boldsymbol{\eta}^{(i)}}^2\Big)^{\frac{1}{2}} \tag{3.7.2}$$

The strange convention above allows one to say that the kinetic energy of a motion t$\to\mathbf{x}(t)$, $t\geq t_0$, of a system of N points, with masses $m_1,\ldots,m_N$ is

$$T(t)=\frac{1}{2}||\dot{\mathbf{x}}(t)||^2 \tag{3.7.3}$$

i.e., it is one-half the square of the velocity of the point representing the system configuration in $\mathcal{R}^{Nd}$ without explicit reference to the masses (which of course are now hidden in the definition of length given by Eq. (3.7.2)).

Let $(U,\boldsymbol{\Xi})$ be a local system of regular coordinates in $\mathcal{R}^{Nd}$ with basis $\Omega\subset\mathcal{R}^{Nd}$ (see Definition 10, p.170) and let $t\to\mathbf{x}(t)$, $t\geq t_0$, be a motion of a system of N point masses, with masses $m_1,\ldots,m_N>0$, taking place for $t\in[t_1,t_2]$ inside U (i.e., $\mathbf{x}(t)\in U$, $\forall\, t\in[t_1,t_2]$).

We can then consider the motion $t\to\mathbf{b}(t)$, $t\in[t-1,t_2]$, "image" in the basis ω of the motion $t\to\mathbf{x}(t)$, $t\in[t_1,t_2]$, via the coordinate system $(U,\boldsymbol{\Xi})$ i.e., the motion such that

$$\mathbf{x}(t)=\boldsymbol{\Xi}(\mathbf{b}(t)),\qquad t\in[t_1,t_2]. \tag{3.7.4}$$

It is obviously possible to express the kinetic energy of the motion $\mathbf{x}$ in terms of the kinematic properties of its image motion $\mathbf{b}$. In fact, if $\boldsymbol{\Xi} = (\boldsymbol{\Xi}^{(1)}, \dots, \boldsymbol{\Xi}^{(N)})$, differentiating Eq. (3.7.4) with respect to t

$$\dot{\mathbf{x}}^{(j)}(t) = \sum_{\ell=1}^{Nd} \frac{\partial \boldsymbol{\Xi}^{(j)}}{\partial \beta_\ell}(\mathbf{b}(t))\dot{b}_\ell(t), \qquad j = 1, \dots, N \tag{3.7.5}$$

where $\mathbf{x}(t) = (\mathbf{x}^{(1)}(t), \dots, \mathbf{x}^{(N)}(t))$, $\mathbf{b}(t) = (b_1(t), \dots, b_{Nd}(t))$, with the usual notations. Hence

$$\begin{aligned} T(t) &= \sum_{j=1}^{N} \frac{m_j}{2} \dot{\mathbf{x}}^(j)^2 \\ &= \sum_{\ell',\ell''=1}^{Nd} \sum_{j=1}^{N} \frac{m_j}{2} \Big(\frac{\partial \boldsymbol{\Xi}^{(j)}}{\partial \beta_{\ell'}}(\mathbf{b}(t)) \frac{\partial \boldsymbol{\Xi}^{(j)}}{\partial \beta_{\ell''}}(\mathbf{b}(t)) \Big) \dot{b}_{\ell'}(t)\dot{b}_{\ell''}(t) \end{aligned} \tag{3.7.6}$$

which will be written as

$$T(t) = \frac{1}{2} \sum_{\ell',\ell''=1} g_{\ell'\,\ell''}(\mathbf{b}(t))\dot{b}_{\ell'}(t)\dot{b}_{\ell''}(t) \tag{3.7.7}$$

having set, $\forall\, \boldsymbol{\beta} \in \Omega$, $\forall\, \ell', \ell'' = 1, \dots, Nd$:

$$g_{\ell'\,\ell''}(\boldsymbol{\beta}) = \sum_{j=1}^{N} m_j \frac{\partial \boldsymbol{\Xi}^{(j)}}{\partial \beta_{\ell'}}(\boldsymbol{\beta}) \cdot \frac{\partial \boldsymbol{\Xi}^{(j)}}{\partial \beta_{\ell''}}(\boldsymbol{\beta}) = g_{\ell''\,\ell'}(\boldsymbol{\beta}) \tag{3.7.8}$$

It is convenient to establish a general definition in connection with Eqs. (3.7.7) and (3.7.8) because of the generality of Eq. (3.7.7) itself.

11 Definition. *The function $\boldsymbol{\beta} \to g(\boldsymbol{\beta})$ defined by Eq. (3.7.8) on Ω and with values in the $Nd \times Nd$ matrices will be called the "kinetic matrix" for the scalar product of Eq. (3.7.1) or, equivalently, for a system of N points in $\mathcal{R}^d$, with masses $m_1, \dots, m_N > 0$, "relative to the local system of regular coordinates $(U, \boldsymbol{\Xi})$ in $\mathcal{R}^{Nd}$".*

Observations. (1) Via Eq. (3.7.7), the kinetic matrix allows one to compute the kinetic energy in arbitrary local coordinates; hence, its name.
(2) Some of the properties of the kinetic matrix will be listed and discussed at the end of the section. For the moment, note that $g(\boldsymbol{\beta})$ is a symmetric matrix whose elements, thought of as functions on Ω, are in $C^\infty(\Omega)$.

For the study of the kinematics of quasi-constrained systems, the following purely geometrical definition is useful.

12 Definition. *Let Σ be a regular surface in $\mathcal{R}^{Nd}$ with codimension s and let U be a neighborhood of a point $\boldsymbol{\xi}_0 \in \Sigma$ on which a system $(U, \boldsymbol{\Xi})$ of local regular coordinates, with basis Ω, is defined.*

Assume that the coordinate system is "adapted" to Σ: (see Definition 10, p.170) i.e., that $\Sigma \cap U$ is described in $(U, \boldsymbol{\Xi})$ by

$$\beta_1 = \ldots = \beta_s = 0 \tag{3.7.9}$$

(a) We shall say that $(U, \boldsymbol{\Xi})$ is "well adapted" to Σ if the kinetic matrix, for the scalar product of Eq. (3.7.1), associated with $(U, \boldsymbol{\Xi})$ has the first principal submatrix which is constant on the plane of Eq. (3.7.9); i.e.. if for $\boldsymbol{\beta} = (0, \ldots, 0, \beta_{s+1}, \ldots, \beta_{Nd}) \in \Omega$ it is, $\forall\, \ell', \ell'' = 1, \ldots, s$:

$$g_{\ell'\,\ell''}(\boldsymbol{\beta}) = g_{\ell'\,\ell''}(0, \ldots, 0, \beta_{s+1}, \ldots, \beta_{Nd}) = \gamma_{\ell', \ell''} \tag{3.7.10}$$

where γ is a $s \times s$ $\boldsymbol{\beta}$-independent matrix.
(b) We shall say that $(U, \boldsymbol{\Xi})$ is "orthogonal" on Σ with respect to the scalar product of Eq. (3.7.1) if $\forall\, \boldsymbol{\beta} = (0, \ldots, 0, \beta_{s+1}, \ldots, \beta_{Nd})$:

$$g_{\ell k}(\boldsymbol{\beta}) = 0, \qquad \ell = 1, 2, \ldots, s;\ k = s+1, \ldots, Nd. \tag{3.7.11}$$

Observations. (1) Let $t \to \mathbf{x}(t)$, $t \in \mathcal{R}_+$, be a motion of N points in $\mathcal{R}^d$, with masses $m_1, \ldots, m_N > 0$, which at some time $\bar{t}$ happens to be in Σ with velocity $\dot{\mathbf{x}}(t)$ "purely transversal" to Σ in the coordinate system $(U, \boldsymbol{\Xi})$, i.e., such that the motion $\mathbf{b}$, image of $\mathbf{x}$ in Ω for t close to $\bar{t}$, has velocity $\dot{\mathbf{b}}(t)$ with components vanishing "along Σ":

$$\dot{\mathbf{b}}(\bar{t}) = (\dot{b}_1(\bar{t}), \ldots, \dot{b}_s(\bar{t}), 0, \ldots, 0). \tag{3.7.12}$$

If the coordinates system is well adapted, then the kinetic energy of $\mathbf{x}$ at time t depends only upon $\dot{\mathbf{b}}(\bar{t})$ but not on the particular position $\mathbf{b}(\bar{t})$.
(2) If the coordinate system $(U, \boldsymbol{\Xi})$ is orthogonal on Σ, the kinetic energy $T(t)$ of a motion $t \to \mathbf{x}(t)$ which for $t = \bar{t}$ crosses Σ is in this instant a sum of two terms: one depending only on $\dot{b}_1(\bar{t}), \ldots, \dot{b}_s(\bar{t})$ and on $\mathbf{b}(\bar{t})$, and the other depending only on $\dot{b}_{s+1}(\bar{t}), \ldots, \dot{b}_{Nd}(\bar{t})$ and on $\mathbf{b}(\bar{t})$:

$$T_1(\bar{t}) = \sum_{\ell', \ell''}^{1,s} g_{\ell'\,\ell''}(\mathbf{b}(\bar{t}))\dot{\beta}_{\ell'}(\bar{t})\dot{b}_{\ell''}(\bar{t}), \quad T_2(\bar{t}) = \sum_{\ell', \ell''}^{s+1,Nd} g_{\ell'\,\ell''}(\mathbf{b}(\bar{t}))\dot{\beta}_{\ell'}(\bar{t})\dot{b}_{\ell''}(\bar{t}), \tag{3.7.13}$$

and $T(\bar{t}) = T_1(\bar{t}) + T_2(\bar{t})$. In other words, one can say that, in such a system of coordinates, for $t = \bar{t}$ the kinetic energy is the sum of the kinetic energies of the component of motion orthogonal to Σ and of the component parallel to it.
(3) Thinking of this, it should become geometrically evident that if Σ is a regular surface in $\mathcal{R}^{Nd}$ and $\boldsymbol{\xi}_0 \in \Sigma$, it will always be possible to construct a system of local coordinates in a neighborhood U of $\boldsymbol{\xi}_0$ which is well adapted and orthogonal to Σ (see Proposition 12 to follow).

The equations of motion can immediately be written in an arbitrary coordinate system, either in the absence or in the presence of constraints, using the following propositions.

9 Proposition. *Let $V \in C^\infty(\mathcal{R}^{Nd})$ be a real-valued function bounded from below. Let $t \to \mathbf{x}(t)$, $t \ge t_0$ be a motion of N points, with masses $m_1, \ldots, m_N > 0$, in $\mathcal{R}^d$ developing under the influence of the force $\mathbf{F}$ with potential energy V. Suppose that for $t \in [t_1, t_2]$, the motion $\mathbf{x}$ takes place in a neighborhood $U \subset \mathcal{R}^{Nd}$ where a local system of regular coordinates $(U, \boldsymbol{\Xi})$ is established with basis $\Omega \subset \mathcal{R}^{Nd}$. Call $\mathbf{b}$ the motion in Ω, image of the considered motion $\mathbf{x}$, for $t \in [t_1, t_2]$, via the coordinate transformation $\boldsymbol{\Xi}$.*
Then $\mathbf{b}$ verifies Lagrange's equations[7] *associated with the Lagrangian:*

$$(\boldsymbol{\alpha}, \boldsymbol{\beta}) \to \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{\ell', \ell''=1}^{Nd} \frac{1}{2} g_{\ell' \, \ell''}(\boldsymbol{\beta}) \alpha_{\ell'} \, \alpha_{\ell''} - V(\boldsymbol{\Xi}(\boldsymbol{\beta})), \tag{3.7.14}$$

where $\boldsymbol{\alpha} \in \mathcal{R}^{Nd}$, $\boldsymbol{\beta} \in \Omega$ and g is the kinetic matrix of Eq. (3.7.1) relative to $(U, \boldsymbol{\Xi})$. Explicitly, such equations are, $\forall \, \ell = 1, \ldots, Nd$:

$$\begin{aligned} \frac{d}{dt}\Big(\sum_{\ell'=1}^{Nd} g_{\ell, \ell'}(\mathbf{b}(t)) \dot{b}_{\ell'}(t)\Big) &= -\Big(\frac{\partial V(\boldsymbol{\Xi}(\boldsymbol{\beta}))}{\partial \beta_\ell}\Big)_{\boldsymbol{\beta} = \mathbf{b}(t)} \\ &+ \frac{1}{2} \sum_{\ell', \ell''=1}^{Nd} \frac{\partial g_{\ell' \, \ell''}}{\partial \beta_\ell}(\boldsymbol{\beta}(t)) \dot{\beta}_{\ell'}(t) \, \dot{\beta}_{\ell''}(t). \end{aligned} \tag{3.7.15}$$

Proof. This is an exercise based on the definition of Lagrangian equations and on the least-action principle as in Proposition 4, §3.3. It will be left to the reader.

10 Proposition. *Consider N points in $\mathcal{R}^d$, with masses $m_1, \ldots, m_N > 0$, and let $\mathbf{F}^{(a)}$ be a conservative force with potential energy $V^{(a)}$, bounded below. Let $\Sigma \subset \mathcal{R}^{Nd}$ be a codimension-s regular surface and suppose that Σ is the set of points $\boldsymbol{\xi} \in R^{Nd}$ such that*

$$\varphi^{(1)}(\boldsymbol{\xi}) = \ldots = \varphi^{(s')}(\boldsymbol{\xi}) = 0, \tag{3.7.16}$$

where $\varphi^{(1)}, \ldots, \varphi^{(s')} \in C^\infty(\mathcal{R}^{Nd})$.
Let $t \to \mathbf{x}(t) \in \Sigma$, $t \in [t_1, t_2]$, be a motion developing in a neighborhood U

[7] If $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \to \widetilde{\mathcal{L}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is a real C^∞ function defined on an open set $W \subset \mathcal{R}^{2M}$, we shall say that a C^∞ function $t \to \mathbf{b}(t)$, $t \in [t_1, t_2]$, such that $(\dot{\mathbf{b}}(t), \mathbf{b}(t)) \in W$, $\forall t \in [t_1, t_2]$, verifies Lagrange's equations associated with $\widetilde{\mathcal{L}}$ if

$$\frac{d}{dt}\left(\frac{\partial \widetilde{\mathcal{L}}}{\partial \alpha_i}(\dot{\mathbf{b}}(t), \mathbf{b}(t))\right) = \frac{\partial \widetilde{\mathcal{L}}}{\partial \beta_i}(\dot{\mathbf{b}}(t), \mathbf{b}(t)), \qquad i = 1, 2, \ldots, M$$

even when W does not coincide with $\mathcal{R}^{2M}$, as usually supposed so far in connection with the Lagrange equations.

of $\boldsymbol{\xi}_0 \in \Sigma$ *under the influence of the force* $\mathbf{F}^{(a)}$ *and of the holonomous ideal, constraint* $\varphi^{(1)}, \ldots, \varphi^{(s')}$ *in the sense of Definition 8, §3.5, p.160. Suppose that* $(U, \boldsymbol{\Xi})$ *is a local system of regular coordinates adapted to* Σ *and with basis* Ω *(see Definition 10, §3.6) and let* $\mathbf{b}$ *be the image motion of* $\mathbf{x}$ *on* Ω*:*

$$\mathbf{b}(t) = \boldsymbol{\Xi}^{-1}(\mathbf{x}(t)), \qquad t \in [t!, t_2]. \tag{3.7.17}$$

Since $\mathbf{b}$ *respects the constraints it is such that*

$$b_1(t) = \ldots = b_s(t) = 0, \qquad t \in [t_1, t_2]. \tag{3.7.18}$$

Then, setting $\mathbf{b}(t) = (0, \ldots .0, \beta_{s+1}(t), \ldots, \beta_{Nd}(t)) \equiv (\mathbf{0}, \mathbf{b}^{(s)}(t))$*, the motion* $t \to \mathbf{b}^{(s)}(t)$ *verifies the Lagrange equations*[8] *associated with a Lagrangian* $\mathcal{L}_0$ *on* $\mathcal{R}^{Nd-s} \times \Omega^{(s)}$ *where* $\Omega^{(s)}$ *denotes the set of points* $\boldsymbol{\beta} = (\beta_{s+1}, \ldots, \beta_{Nd}) \in \mathcal{R}^{Nd-s}$ *such that* $(\mathbf{0}, \boldsymbol{\beta}) \in \Omega$ *(*$\mathbf{0}$ *being the origin in* $\mathcal{R}^s$*). The Lagrangian* $\mathcal{L}_0$ *is defined by*

$$\mathcal{L}_0(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{2} \sum_{\ell', \ell''}^{s+1, Nd} g_{\ell', \ell''}(\mathbf{0}, \boldsymbol{\beta})\, \alpha_{\ell'} \alpha_{\ell''} - V^{(a)}(\boldsymbol{\Xi}(\mathbf{0}, \boldsymbol{\beta})), \tag{3.7.19}$$

$\forall\, (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}^{Nd-s} \times \Omega^{(s)}$ *(*$\mathbf{0}$ *being the origin in* $\mathcal{R}^s$*). Hence, the equations of motion for* $\boldsymbol{\beta}^{(s)}(t)$ *are,* $\forall\, \ell = s-1, \ldots, Nd$

$$\begin{aligned} &\frac{d}{dt}\Big(\sum_{\ell'=1}^{Nd} g_{\ell,\ell'}(\mathbf{0}, \mathbf{b}^{(s)}(t)) \dot{b}^{(s)}_{\ell'}(t) \Big) = -\Big(\frac{\partial V(\boldsymbol{\Xi}(\mathbf{0}, \boldsymbol{\beta}))}{\partial \beta_\ell} \Big)_{\boldsymbol{\beta} = \mathbf{b}^{(s)}(t)} \\ &+ \frac{1}{2} \sum_{\ell', \ell''=1}^{Nd} \frac{\partial g_{\ell'\,\ell''}(\mathbf{0}, \mathbf{b}^{(s)}(t))}{\partial \beta_\ell} \dot{\beta}^{(s)}_{\ell'}(t)\, \dot{\beta}^{(s)}_{\ell''}(t). \end{aligned} \tag{3.7.20}$$

PROOF. Proposition 10 can be proved as a corollary to Proposition 8, §3.5, p.160, and it will be left to the reader as an important exercise.

Through the Propositions 9 and 10 and the above definitions, it is now possible to reformulate and make precise the Problem (1) posed at the end of §3.6, p.174. It appears to be naturally related to the following definition.

13 Definition. *Let* (Σ, W, λ) *be a model for a bilateral conservative constraint for a system of* N *points in* $\mathcal{R}^d$*, with masses* $m_1, \ldots, m_N$*. Suppose that* Σ *has codimension* s *and that it is described by Eq. (3.7.16).*
Let $V^{(a)}$ *be a real-valued* $C^\infty(\mathcal{R}^{Nd})$ *function bounded below and let* $t \to \mathbf{x}_\lambda(t)$*,* $t \in \mathcal{R}_+$*, be the motion of the system developing under the influence of the force with potential energy*

[8] See footnote 7.

$$\boldsymbol{\xi} \to V(\boldsymbol{\xi}) = V^{(a)}(\boldsymbol{\xi}) + \lambda\, W(\boldsymbol{\xi}) \tag{3.7.21}$$

following the initial datum

$$\mathbf{x}_\lambda(0) = \boldsymbol{\xi}_0 \in \Sigma, \qquad \dot{\mathbf{x}}(\mathbf{0}) = \boldsymbol{\eta}_0 \in \mathcal{R}^{Bd}. \tag{3.7.22}$$

We shall say that (Σ, W, λ) is a model of an "ideal approximate constraint" if, $\forall\, V^{(a)}, \boldsymbol{\xi}_0, \boldsymbol{\eta}_0$ as above:
(i) The following limit exists:

$$\lim_{\lambda \to +\infty} \mathbf{x}_\lambda(t) = \mathbf{x}(t), \qquad \forall\, t \geq 0 \tag{3.7.23}$$

(ii) The function $t \to \mathbf{x}(t)$ is a motion developing on Σ: under the influence of the active force $\mathbf{F}^{(a)}$, with potential energy $V^{(a)}$, and of the ideal constraint $\varphi^{(1)}, \ldots, \varphi^{(s')}$ to Σ, in the sense of Definition 8, §3.5.
(iii) The initial datum verified by $\mathbf{x}$ is

$$\mathbf{x}(0) = \boldsymbol{\xi}_0, \qquad \dot{\mathbf{x}}(0) = \boldsymbol{\eta}_0^{\Sigma}, \tag{3.7.24}$$

where $\boldsymbol{\eta}_0^{\Sigma}$ is the orthogonal projection of $\boldsymbol{\eta}_0$ on the tangent plane to Σ in $\boldsymbol{\xi}_0$, with respect to the scalar product of Eq. (3.7.1) [see Observation (1) below].

Observations. (1) Let $(U, \boldsymbol{\Xi})$ be a local system of regular coordinates with basis Ω. Let $\boldsymbol{\xi}_0 \in U$ and suppose that $(U, \boldsymbol{\Xi})$ is adapted and orthogonal on Σ: it will be, by Eq. (3.7.5),

$$\boldsymbol{\eta}_0 = \sum_{i=1}^{Nd} \frac{\partial \boldsymbol{\Xi}}{\partial \beta_i}(\boldsymbol{\beta}_0)\, \alpha_i^0 \tag{3.7.25}$$

where $\boldsymbol{\beta}_0$ are the coordinates of $\boldsymbol{\xi}_0$ in $(U, \boldsymbol{\Xi})$ and $\boldsymbol{\alpha}^0 \in \mathcal{R}^{Nd}$ is a suitable vector. Then the projection $\boldsymbol{\eta}_0^{\Sigma}$ is, by definition,

$$\boldsymbol{\eta}_0^{\Sigma} = \sum_{i=s+1}^{Nd} \frac{\partial \boldsymbol{\Xi}}{\partial \beta_i}(\boldsymbol{\beta}_0)\, \alpha_i^0 \tag{3.7.26}$$

It could be checked that $\boldsymbol{\eta}_0^{\Sigma}$ does not depend on the coordinate system $(U, \boldsymbol{\Xi})$ provided the latter is adapted and orthogonal on Σ.
(2) By the above definition, if (Σ, W, λ) is a model of an approximate ideal constraint, the motions of the system starting on Σ and developing under the influence of a conservative force with potential energy $V^{(a)} + \lambda W$ are, for large λ, well approximated by the motions of the same system subject to s' ideal constraints $\varphi^{(1)}, \ldots, \varphi^{(s')}$, determining Σ, and to an active force with potential energy $V^{(a)}$, in the sense of Definition 8,§3.5, p.160. Of course, it would be desirable, and necessary in order to make quantitative statements, to have estimates of the difference between $\mathbf{x}(t)$ and $\mathbf{x}_\lambda(t)$ in terms of λ: this will be done, in some cases, in§3.8.

(3) Problem 1, p.174, §3.6, is equivalent to the following question: given (Σ, W, λ), how does one recognize whether this is a model of an approximate ideal constraint?
(4) Assume that for $t \in [t_1, t_2]$, the motion $t \to \mathbf{x}_\lambda(t)$ takes place for large enough λ in a neighborhood U such that $U \cap \Sigma \neq \emptyset$ and suppose that a local system of regular coordinates is established on U: $(U, \boldsymbol{\Xi})$ with basis Ω adapted to Σ. Let $t \to \mathbf{b}_\lambda(t)$, $t \in [t_1, t_2]$ be the image in Ω of $\mathbf{x}_\lambda$ considered for $t \in [t_1, t_2]$. Items (i) and (ii) of Definition 13 are equivalent to:
(i_1) There is a limit

$$\lim_{\lambda \to +\infty} \mathbf{b}_\lambda(t) = \mathbf{b}(t) \equiv (\mathbf{0}, \mathbf{b}^{(s)}(t)), \qquad t \in [t_{1,2}]. \tag{3.7.27}$$

where $\mathbf{0}$ is the origin in $\mathcal{R}^s$ and $\mathbf{b}^{(s)}(t)$ is a suitable $\mathcal{R}^{Nd-s}$-valued function.
(ii_1) t→ $\mathbf{b}(t)$ is a $C^\infty([t_1, t_2])$ function verifying Eqs. (3.7.18) and (3.7.20), $\forall\, t \in [t_1, t_2]$.

Condition (iii) is equivalent to:
(iii_1) If $\boldsymbol{\xi}_0 \in U$ and $(U, \boldsymbol{\Xi})$ is orthogonal on Σ:

$$\boldsymbol{\Xi}(\mathbf{b}(0)) = \boldsymbol{\xi}_0, \quad \dot{b}_i(0) = 0, \qquad i = 1, 2, \ldots, s. \tag{3.7.28}$$

In the next section we shall discuss an important sufficient perfection criterion for approximate conservative constraints and the perfection will checked in the form of (i_1), (ii_1), and (iii_1) above.

We conclude this section by stating some simple properties of the kinetic matrices on $\mathcal{R}^{Nd}$ associated with the scalar product of Eq. (3.7.1) in a local system of regular coordinates. We shall also sketch the proof of some metric properties of the regular surfaces (i.e., the existence of well-adapted and orthogonal coordinates).

11 Proposition. *The matrix $g(\boldsymbol{\beta})$ defined in Eq. (3.7.6) on Ω is, $\forall\, \boldsymbol{\beta} \in \Omega$, symmetric and positive definite. The matrix elements of $g(\boldsymbol{\beta})$, as well as the matrix elements of the matrices inverting $g(\boldsymbol{\beta})$ or any of its principal submatrices, are in $C^\infty(\Omega)$ as functions of $\boldsymbol{\beta} \in \Omega$.*

There exists a positive continuous function $\boldsymbol{\beta} \to C(\boldsymbol{\beta})$, defined on Ω, such that if $\mu(\boldsymbol{\beta})$ is a $q \times q$ principal submatrix of $g(\boldsymbol{\beta})$ or an inverse to such a matrix, then $\forall\, \boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_q) \in \mathcal{R}^q$:

$$C(\boldsymbol{\beta}) \sum_{i=1}^{q} \sigma_i^2 \leq \sum_{\ell', \ell''}^{1,q} \mu_{\ell'\,\ell''}(\boldsymbol{\beta}) \sigma_{\ell'} \sigma_{\ell''} \leq C(\boldsymbol{\beta})^{-1} \sum_{i=1}^{q} \sigma_i^2. \tag{3.7.29}$$

Observation. We recall that a $q \times q$ matrix μ is called positive definite if it is symmetric and

$$\sum_{\ell',\ell''}^{1,q} \mu_{\ell'\,\ell''}\sigma_{\ell'}\sigma_{\ell''} > 0, \qquad \forall\,\boldsymbol{\sigma} \in \mathcal{R}^q,\ \boldsymbol{\sigma} \neq \mathbf{0} \tag{3.7.30}$$

(see Appendix F).

PROOF. The symmetry is obvious and has already been remarked on [see Eq. (3.7.8)]. The positivity of $g(\boldsymbol{\beta})$ follows from its kinematic interpretation: in fact, $\frac{1}{2}\sum_{\ell',\ell''}^{1,Nd} g_{\ell'\,\ell''}(\boldsymbol{\beta})\sigma_{\ell'}\sigma_{\ell''}$ is the kinetic energy of a motion which at some time happens to be in $\boldsymbol{\Xi}(\boldsymbol{\beta})$ with velocity $\dot{\mathbf{x}}^{(j)} = \sum_{\ell=1}^{Nd} \frac{\partial \boldsymbol{\Xi}^{(j)}}{\partial \beta_\ell}\,\sigma_\ell$, $j = 1,\ldots,N$ [see Eq. (3.7.5)].

If $\sigma \neq 0$, then $\dot{\mathbf{x}} \neq \mathbf{0}$ because the coordinate system is regular [see Eq. (3.6.10)]. Hence, $\sum_{\ell',\ell''}^{1,Nd} g_{\ell'\,\ell''}(\boldsymbol{\beta})\sigma_{\ell'}\sigma_{\ell''} = \sum_{j=1}^{N} m_j\,(\dot{\mathbf{x}}^{(j)})^2 > 0$.

From algebra, it is known that a matrix $g(\boldsymbol{\beta})$ is positive definite if and only if all its principal submatrices are positive definite: in such case also their inverse matrices are all positive definite and all the mentioned matrices have a positive determinant. Furthermore, if μ is a $q \times q$ positive definite matrix, there is a positive continuous function of its matrix elements such that

$$C_1(\mu)\sum_{i=1}^{q}\sigma_i^2 \le \sum_{\ell',\ell''}^{1,q} \mu_{\ell'\,\ell''}\sigma_{\ell'}\sigma_{\ell''} \le C_1(\mu)^{-1}\sum_{i=1}^{q}\sigma_i^2 \tag{3.7.31}$$

(see, also, Appendix F, Corollary 3 and related exercises).

Then the proposition is a consequence of these algebraic properties and of the observation that all the mentioned matrix elements are in $C^\infty(\Omega)$: in fact they are obtained by taking products and sums of matrix elements of $g(\boldsymbol{\beta})$ and possibly dividing the results by products of determinants of some principal submatrices of the matrix $g(\boldsymbol{\beta})$, which are in turn positive by what has just been mentioned. mbe

The following proposition concerns the existence of a local system of regular coordinates $(U, \boldsymbol{\Xi})$ well adapted and orthogonal to a regular surface Σ in the vicinity of one of its points $\boldsymbol{\xi}_0$.

12 Proposition. *Given a regular surface $\Sigma \subset \mathcal{R}^{Nd}$ with codimension s and given the scalar product of Eq. (3.7.1) and $\boldsymbol{\xi}_0 \in \Sigma$, it is always possible to find a neighborhood U of $\boldsymbol{\xi}_0$ on which it is possible to define a local regular system of coordinates well adapted and orthogonal to Σ. It is even possible to construct it so that the kinetic matrix is, in the basis points corresponding to $\Sigma \cap U$,*

$$g_{\ell\,\ell'} = \gamma\,\delta_{\ell\,\ell'}, \qquad \ell,\ell' = 1,2,\ldots,s, \quad \gamma > 0 \tag{3.7.32}$$

("Fermi coordinates" on Σ).

PROOF. Only a sketch will be given, leaving to the reader the task of completing the proof (see, also, exercises and problems at the end of this section).

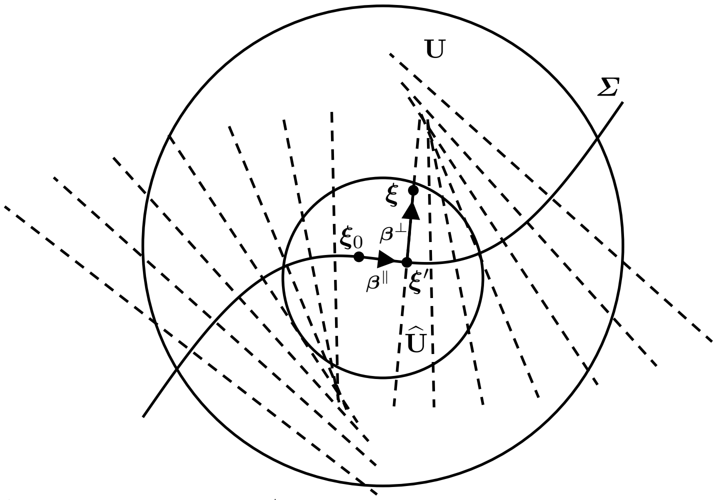


Figure3.1

Fig.3.1. ξ has coordinates $(\beta'', \beta^{\perp})$ and $\beta'' = \{$abscissa of ξ' on $\Sigma\}$.

A first reading should be intended just in order to get some ideas about the proof and the geometrical meaning of Proposition 12: completing the details should then become easier and transparent.

Let $\boldsymbol{\xi}_0 \in \Sigma$ and let U be a bounded neighborhood on which a local system of regular coordinates adapted to Σ, $(U, \boldsymbol{\Xi})$, is established. Suppose that $\boldsymbol{\Xi}^{-1}(\boldsymbol{\xi}_0) = \mathbf{0}$, say. An orthogonal and well-adapted system $(\widehat{U}, \widehat{\boldsymbol{\Xi}})$ will be built by suitably choosing $\widehat{U} \subset U$. The system is geometrically illustrated in the case $d = 2, s = 1$, in Fig.3.1.

The construction proceeds as follows: at every point $\boldsymbol{\xi}' \in \Sigma \cap U$, consider a hyperplane in $\mathcal{R}^{Nd}$ orthogonal to Σ in $\boldsymbol{\xi}'$ in the sense of the orthogonality associated with the scalar product of Eq. (3.7.1). Denote this hyperplane by $\pi(\boldsymbol{\xi}')$ (dotted lines in Fig. 3.1). Fix on $\pi(\boldsymbol{\xi}')$ a system of Cartesian mutually orthogonal axes with unit vectors $\mathbf{e}_1(\boldsymbol{\xi}'), \ldots, \mathbf{e}_s(\boldsymbol{\xi}')$, the orthogonality being in the sense of the scalar product of Eq. (3.7.1) and the length of the axes being measured in the same sense.

Choose the above unit vectors so that the points $\boldsymbol{\xi}' + \mathbf{e}_i(\boldsymbol{\xi}')$, $i = 1, \ldots, s$, have coordinates which are C^∞ functions of the $Nd - s$ nontrivial coordinates of $\boldsymbol{\xi}'$ in (U, Σ); i.e, choose the Cartesian axes "so that they are C^∞ functions of $\boldsymbol{\xi}' \in \Sigma \cap U$".

There is a neighborhood U' of $\boldsymbol{\xi}_0$, $U' \subset U$ such that every point $\boldsymbol{\xi} \in U'$ is on a unique plane $\pi(\boldsymbol{\xi}')$ with $\boldsymbol{\xi}'$ suitably chosen on $\Sigma \cap U$ ("consequence of the finite curvature of Σ").

To every $\boldsymbol{\xi} \in U'$, we then associate Nd coordinates $\widehat{\boldsymbol{\beta}}$: the first s of them, denoted $\widehat{\boldsymbol{\beta}}^{\perp} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_s)$, are the coordinates of $\boldsymbol{\xi}$ in the Cartesian frame chosen on the plane $\pi(\boldsymbol{\xi}')$ containing $\boldsymbol{\xi}$; the remaining $Nd - s$ coordinates $\widehat{\boldsymbol{\beta}}^{\parallel} = (\widehat{\beta}_{s+1}, \ldots, \widehat{\beta}_{Nd})$ are the coordinates with $\widehat{\beta}_i = \beta_i, i = s+1, \ldots, Nd$, if $\boldsymbol{\xi}'$ has in $(U, \boldsymbol{\Xi})$ coordinates $(0, \ldots, 0, \beta_{s+1}, \ldots, \beta_{Nd})$.

Setting $\boldsymbol{\xi} = \widehat{\boldsymbol{\Xi}}(\widehat{\boldsymbol{\beta}})$, one can check that as $\boldsymbol{\xi}$ varies in U', the point $\widehat{\boldsymbol{\beta}}$ varies in some open set $\Omega' \in \mathcal{R}^{Nd}$. Furthermore, the function $\widehat{\boldsymbol{\Xi}}$ is a C^∞-invertible map of Ω' onto U' which is not singular, i.e., its Jacobian matrix [see Eq. (3.6.10)] has a never-vanishing determinant if U' is small enough. Let $\widehat{S}^{Nd}$ be a sphere contained in Ω' centered at $\boldsymbol{\Xi}^{-1}(\boldsymbol{\xi}_0) = \mathbf{0}$ and set $\widehat{U} = \widehat{\boldsymbol{\Xi}}(\widehat{S}^{Nd})$. Then, essentially by construction, the pair $(\widehat{U}, \widehat{\boldsymbol{\Xi}})$ is a coordinate system which is of Fermi type on Σ with basis $\widehat{S}^{Nd}$ and $\gamma = 1$.

The difficulty, in a rigorous proof, lies in the justification of the actual possibility of the various "choices" involved in the above descriptive argument, and in checking the validity of the statements claimed about the uniqueness of the plane $\pi(\boldsymbol{\xi}')$ through $\boldsymbol{\xi}'$ and on the non singularity of the Jacobian matrix

$$\widehat{J}(\widehat{\boldsymbol{\beta}}) = \frac{\partial \widehat{\boldsymbol{\xi}}}{\partial \widehat{\boldsymbol{\beta}}}(\widehat{\boldsymbol{\beta}}), \qquad \widehat{\boldsymbol{\beta}} \in \widehat{S}^{Nd}. \tag{3.7.33}$$

The main idea is to use the implicit function theorem to check the above properties, (see the following problems for some more details.)

3.7.1 Exercises and Problems

1. Establish an orthogonal regular system of coordinates well adapted to the circle $\Gamma \subset \mathcal{R}^2$, with radius 1 and center at the origin, with respect to the scalar product $\boldsymbol{\eta}\cdot\boldsymbol{\chi} = \eta_1\chi_1 + \eta_2\chi_2$, in the neighborhood of a generic point $\xi_0 \in \Gamma$.

2. Same as Problem 1, replacing Γ with the parabola $y = x^2$, the hyperbola $xy = 1$, or the ellipse $x^2 2/a^2 + y^2/b^2 = 1, a, b > 0$.

3.* Let $\Gamma \subset \mathcal{R}^2$ be a simple C^∞ curve in $\mathcal{R}^2$ parameterized in terms of its curvilinear abscissa $s \in \mathcal{R}$ as:

$$\begin{cases} \xi_1 = X_1(s), & (X_1'(s))^2 + (X_2'(s))^2 = 1, \\ \xi_2 = X_2(s), & \lim_{s\to\pm\infty} (X_1(s))^2 + (X_2(s))^2 = +\infty, \end{cases}$$

where X_1', X_2' are the derivatives of X_1, X_2. For every point on Γ with abscissa $s \in \mathcal{R}$, consider the normal line $\mathbf{n}(s)$ with equations $\xi_1 X_1'(s) + \xi_2 X_2'(s) = 0$. Show that given $R > 0$, there is $\delta > 0$ such that the segments of length 2δ cut around $(X_1(s), X_2(s))$ on the line $\mathbf{n}(s)$ are pairwise disjoint, $\forall\, 0 \le |s| < R$. (*Hint:* Define for $|s| < R$, $|\sigma| < \delta$:

$$F_1(s,\sigma) = X_1(s) + \sigma\, X_2'(s), \qquad F_2(s,\sigma) = X_2(s) - \sigma\, X_1'(s)$$

and note that the equations $F_1(s,\sigma) = F_1(\overline{s},\overline{\sigma}), F_2(s,\sigma) = F_2(\overline{s},\overline{\sigma})$ thought of as equations for (s,σ) parameterized by $\overline{s},\overline{\sigma}$ have $s = \overline{s}, \sigma = \overline{\sigma}$ as a unique solution near $\overline{s},\overline{\sigma}$, if $\overline{\sigma}$ is small (using the implicit function theorem). Then, by using the possibility of choosing δ small, show that they have a unique solution in the entire region $|s| < R, |\sigma| < \delta$, etc.)

4.* In the context of Problem 3, show that there is $\delta' < \delta$ such that the image via (F_1, F_2) of $(-R,R) \times (-\delta', \delta')$ is a neighborhood U of $(F_1(0,0), F_2(0,0)) \in \Gamma$. where the map $(s,\sigma) \longleftrightarrow (F_1(s,\sigma), F_2(s,\sigma))$ is invertible, C^∞ and with nonvanishing Jacobian, i.e., the map $(U, \mathbf{F})$ is an adapted system of local regular coordinates.

5.* In the context of Problem 4, compute the kinetic matrix $g(s,\sigma)$ and show that $(U, \mathbf{F})$ is a well-adapted orthogonal coordinate system for Γ, with respect to the scalar product in Problem 1, and g is on Γ a 2×2 diagonal matrix.

6.* Let $z = s + i\sigma$, $(s,\sigma) \in \mathcal{R}^2$, and let f be a function on $\mathcal{C}$ admitting a representation $f(z) = \sum_{n=0}^{\infty} c_n z^n$ with $c_n \xrightarrow[n\to\infty]{} 0$ so fast that the series has an infinite radius of convergence. Also suppose that $f'(z_0) = \sum_{n=0}^{\infty} n c_n z_0^{n-1} \neq 0$ for some $z_0 \in \mathcal{C}$.
Let γ_1, γ_2 be two segments of regular curves crossing at $z_0 \in \mathcal{C}$ and there forming an angle φ_0. Show that the f images of γ_1, and γ_2, $f(\gamma_1)$ and $f(\gamma_2)$, cross at $f(z_0)$ forming the same angle φ_0. (*Hint:* Let $\{dz_1\}$, and $\{dz_2\}$ be two infinitesimal segments in z_0 directed along γ_1, and γ_2. Show that $f(\{dz_1\}) = f'(z_0)\{dz_1\}$, $f(\{dz_2\}) = f'(z_0)\{dz_2\}$, where $f'(z_0) = \sum_{n=0}^{\infty} n c_n z_0^{n-1} \neq 0$, and interpret this as saying that dz_1, and dz_2 are transformed into two infinitesimal segments emerging from $f(z_0)$, elongated by a factor $\varrho_0 = |f'(z_0)|$, and rotated by an angle $\theta_0 = \arg f'(z_0)$ ("conformal mapping property").)

7.* In the context of Problem 6, suppose that $z \to f(z)$ is one to one near z_0 and that $f'(z_0) \neq 0$. Call U a neighborhood of z_0 where this happens. Show that the map $(s,\sigma) \in U \to (s',\sigma') = (\mathcal{R}\mathrm{e}\, f(z), \mathcal{I}\mathrm{m}\, f(z))$ (i.e., $z \to f(z)$) maps U onto a neighborhood V of $(\mathcal{R}\mathrm{e}\, f(z_0), \mathcal{I}\mathrm{m}\, f(z_0))$ and establishes a local system of regular coordinates on V.
Let U be a disk around z_0 and $z_0 = 0$. Consider the curve in V whose equations are $s' = \mathcal{R}\mathrm{e}\, f(s), \sigma' = \mathcal{I}\mathrm{m}\, f(s)$ for $(s,0) \in U$. Show that the above coordinate system is orthogonal on the curve Γ image of the points in U of the form $(s,0)$.

8.* Without use of complex functions, extend the argument of Problem 3 to a regular surface Σ in $\mathcal{R}^d$ with Σ and d arbitrary, using the ordinary scalar product in $\mathcal{R}^d$. (*Hint:* Follow the pattern of the sketch of the proof of Proposition 12 and of the Problem 3, using the implicit function theorem.)

3.8 A Perfection Criterion for Approximate Constraints

This section is devoted to the analysis of the following interesting proposition, "Arnold's theorem", see historical note on p.211.

13 Proposition. *Consider N points, with masses $m_1, \ldots, m_N > 0$, and a model of bilateral conservative approximate constraint (Σ, W, λ) with codimension s. Suppose that $\forall\, \boldsymbol{\xi} \in \Sigma$, there is a neighborhood U admitting a system of local regular coordinates $(U, \boldsymbol{\Xi})$ with basis Ω, well adapted and orthogonal on Σ with respect to the scalar product of Eq. (3.7.1), and such that*

$$W(\boldsymbol{\Xi}(\boldsymbol{\beta})) = \overline{W}(\beta_1, \ldots, \beta_s), \tag{3.8.1}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_s, \beta_{s+1}, \ldots, \beta_{Nd})$ and W is a real $C^\infty(\mathcal{R}^s)$ function, vanishing at the origin and having a strict minimum there.
Then the constraint model (Σ, W, λ) is an ideal approximate constraint, in the sense of Definition 13, p.180.

Observations.
(1) We already noted that it is always possible to find a neighborhood U of $\boldsymbol{\xi}_0$ on which a local system of regular coordinates, well adapted and orthogonal on Σ, can be established. In general, however, the functions $\boldsymbol{\beta} \to W(\boldsymbol{\Xi}(\boldsymbol{\beta}))$

will depend on all the Nd coordinates of $\boldsymbol{\beta} \in \Omega$ and not just on the first s (one can say that W will not, in general, be "purely orthogonal" to the constraint).
(2) Before proceeding to the proof, let us discuss the following example.

Example. Consider a two points system, with masses $m_1, m_2 > 0$, in $\mathcal{R}^3$ and the constraint model ("rigid link at distance ℓ") defined by

$$\Sigma = \{\boldsymbol{\xi} \,|\, \boldsymbol{\xi} = (\boldsymbol{\xi}^{(1)}, \boldsymbol{\xi}^{(2)}),\, |\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}| = \ell \tag{3.8.2}$$

$$W(\boldsymbol{\xi}^{(1)}, \boldsymbol{\xi}^{(2)}) = \Big((\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)})^2 - \ell^2\Big)^2 \tag{3.8.3}$$

where $\ell > 0$ is given. Let us check that the approximate conservative constraint model (Σ, W, λ) verifies the assumptions of Proposition 13. Define the following Baricentric-Polar coordinates:

$$\begin{aligned} \beta_1 =& |\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}| - \ell \equiv \varrho - \ell, \quad \beta_2 = \theta, \qquad \beta_3 = \varphi, \\ \beta_4 =& (\boldsymbol{\xi}_G)_1, \qquad \beta_5 = (\boldsymbol{\xi}_G)_2, \qquad \beta_6 = (\boldsymbol{\xi}_G)_3 \end{aligned} \tag{3.8.4}$$

where $(\varrho, \theta, \varphi)$ are the polar coordinates of the vector $\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}$ and $\boldsymbol{\xi}_B$ is the vector determining the baricenter position:

$$\boldsymbol{\xi}_G = \frac{m_1 \boldsymbol{\xi}^{(1)} + m_2 \boldsymbol{\xi}^{(2)}}{m_1 + m_2} \tag{3.8.5}$$

Through Eq. (3.8.4), one can easily establish a regular local system of coordinates $(U, \boldsymbol{\Xi})$ in the vicinity of any point $\boldsymbol{\xi}_0 \in \Sigma$ such that $\theta \in (0, \pi)$, $\varphi \in (0, 2\pi)$ (which, by the arbitrariness of the choice of Cartesian axes, is not a real restriction). This reference system is adapted to Σ, and Σ is given by $\beta_1 = 0$. To check that it is well adapted and orthogonal, for the scalar product of Eq. (3.7.1), compute the kinetic matrix remarking that

$$m_1 (\dot{\boldsymbol{\xi}}^{(1)})^2 + m_2 (\dot{\boldsymbol{\xi}}^{(2)})^2 = (m_1 + m_2)\, \dot{\boldsymbol{\xi}}_G^2 + \frac{2 m_1 m_2}{m_1 + m_2} (\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)})^2 \tag{3.8.6}$$

which follows immediately from the relations

$$\boldsymbol{\xi}^{(1)} = \boldsymbol{\xi}_G + \frac{m_2}{m_1 + m_2} (\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}), \quad \boldsymbol{\xi}^{(2)} = \boldsymbol{\xi}_G - \frac{m_1}{m_1 + m_2} (\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}), \tag{3.8.7}$$

by differentiation and some algebra. Since

$$(\dot{\boldsymbol{\xi}}^{(1)} - \dot{\boldsymbol{\xi}}^{(2)})^2 = \dot{\varrho}^2 + \varrho^2 \dot{\theta}^2 + \varrho^2 (\sin\theta)^2 \dot{\varphi}^2, \tag{3.8.8}$$

which can be seen by recalling that the line element in polar coordinate is $d\varrho^2 + \varrho^2 d\theta^2 + \varrho^2 (\sin\theta)^2 d\varphi^2$ and that $(\varrho, \theta, \varphi)$ are just the polar coordinates of $\boldsymbol{\xi}^{(1)} - \boldsymbol{\xi}^{(2)}$, it follows that Eqs. (3.8.6) and (3.8.8) give

$$\begin{aligned} m_1(\dot{\boldsymbol{\xi}}^{(1)})^2 + m_2(\dot{\boldsymbol{\xi}}^{(2)})^2 &= (m_1+m_2)(\dot\beta_4^2+\dot\beta_5^2+\dot\beta_6^2)\\ &+\frac{2m_1m_2}{m_1+m_2}(\dot\beta_1^2+(\ell+\beta_1)^2\dot\beta_2^2+(\ell+\beta_1)^2(\sin\beta_2)^2\dot\beta_3^2), \end{aligned} \tag{3.8.9}$$

showing that the coordinates of Eq. (3.8.4) are well adapted and orthogonal on Σ (in fact, $g_{1\,1}(\boldsymbol{\beta}) \equiv \frac{2m_1m_2}{m_1+m_2}$ and the quadratic form of Eq. (3.8.9) does not contain the mixed terms $\dot\beta_1\,\dot\beta_\ell$, $\ell > 1$).

In this coordinate system, the constraint structure function W of Eq. (3.8.3) is simply $(\beta_1^2 + 2\ell\beta_1)^2$, i.e., it depends only on β_1.

A further example is the model (Σ, W, λ) for a single point in $\mathcal{R}^3$ bound to a regular surface σ: $\Sigma = \{\boldsymbol{\xi}\,|\,\boldsymbol{\xi}\in\sigma\}$, and W is a C^∞ function of $\boldsymbol{\xi}$, positive outside Σ and having, for $\boldsymbol{\xi}$ close enough to Σ, the form

$$W(\boldsymbol{\xi}) = (\mathbf{n}\cdot(\boldsymbol{\xi}-\widetilde{\boldsymbol{\xi}}))^2 \tag{3.8.10}$$

where $\widetilde{\boldsymbol{\xi}}$ is the point on σ closest to $\boldsymbol{\xi}$ and $\mathbf{n}$ is a unit vector normal to σ in $\widetilde{\boldsymbol{\xi}}$. (The proof is left as a problem.)

Proof (of Proposition 13). Let $(\boldsymbol{\eta}_0, \boldsymbol{\xi}_0)$ be an initial datum for the given system of point masses, with $\boldsymbol{\xi}_0 \in \Sigma$.

Fix $\lambda \geq 1$ and a function $V^{(a)} \in C^\infty(\mathcal{R}^{Nd})$ bounded from below. Let U a neighborhood of the point $\boldsymbol{\xi}_0$ where it is possible to define a local system of regular coordinates $(U, \boldsymbol{\Xi})$ with basis Ω, well adapted and orthogonal on Σ and such that Eq. (3.8.1) holds in this system. Suppose $\boldsymbol{\Xi}^{-1}(\boldsymbol{\xi}_0) = \mathbf{0}$. We also suppose, for the sake of simplicity, that W has a rather special form:

$$\overline{W}(\beta_1,\ldots,\beta_\ell) = \frac{1}{2}\sum_{i=1}^{s}\beta_i^2 \tag{3.8.11}$$

In spite of the particularity of Eq. (3.8.11), this is an assumption that can be eliminated through some formal complications which would only make the true difficulties of the problem and the solution method more obscure (see problems at the end of this section).

Denote $t \to \mathbf{x}_\lambda(t)$, $t \in \mathcal{R}_+$, the motion that the N points perform under the influence of the force with potential energy $V^{(a)} + \lambda W$ starting from the initial datum $(\boldsymbol{\eta}_0, \boldsymbol{\xi}_0)$. By energy conservation it follows that

$$\sum_{i=1}^{N}\frac{m_i}{2}(\dot{\mathbf{x}}_\lambda(t))^2 + V^{(a)}(\mathbf{x}_\lambda(t)) + \lambda W(\mathbf{x}_\lambda(t)) = E \tag{3.8.12}$$

is a constant in t and

$$E = \sum_{i=1}^{N}\frac{m_i}{2}(\boldsymbol{\eta}_0^{(i)})^2 + V^{(a)}(\boldsymbol{\xi}_0) \tag{3.8.13}$$

is λ independent because $\boldsymbol{\xi}_0 \in \Sigma$ and W vanishes on Σ.

Then Eq. (3.8.12) and the assumed boundedness of $V^{(a)}$ imply that

$$\sum_{i=1}^{N} \frac{m_i}{2}(\dot{\mathbf{x}}_\lambda(t))^2 \le E + \sup(-V^{(a)}(\boldsymbol{\xi})) \stackrel{def}{=} C < +\infty. \tag{3.8.14}$$

If S_ϱ denotes a closed ball with radius ϱ and center $\boldsymbol{\xi}_0$, contained in U, Eq. (3.8.14) will imply that the motion $t \to \mathbf{x}_\lambda(t)$ will develop remaining inside S_ϱ for $t \in [0, T]$, i.e., $\mathbf{x}_\lambda(t) \in S_\varrho$, $\forall t \in [0, T]$, if T is chosen so that

$$T\sqrt{\frac{2C}{\min_j m_j}} < \varrho \tag{3.8.15}$$

Fix T verifying Eq. (3.8.15) consider the motion $t \to \mathbf{x}_\lambda(t)$, $t \in [0, T]$.

Existence of the limit $\lim_{\lambda\to+\infty} \mathbf{x}_\lambda(t) = \mathbf{x}(t)$ and validity of Eqs. (3.7.18) and (3.7.20) will be shown only for $t \in [0, T]$. The treatment of the general case (t arbitrarily large) contains some additional difficulties and will not be discussed in detail. Such difficulties have a geometrical character and depend on the fact that $(U, \boldsymbol{\varXi})$ is generally a local coordinate system and not a global one for all of $\varSigma$, see Problem 5 at the end of this section.

First it will be shown that the motion $t \to \mathbf{x}_\lambda(t)$, $t \in [0, T]$, tends to evolve on $\varSigma$ as $\lambda \to +\infty$.

This is a simple consequence of energy conservation and of the positivity (only) of W: it does not depend on the special hypothesis on the constraint nature [Eq. (3.8.11)], but it would be valid even for general approximate conservative constraints.

Let $t \to \mathbf{b}_\lambda(t)$, $t \in [0, T]$, be the image of the motion $\mathbf{x}_\lambda$, observed for $t \in [0, T]$, in the basis $\varOmega$ of the coordinate system: $\mathbf{b}_\lambda(t) = \boldsymbol{\varXi}^{-1}(\mathbf{x}_\lambda(t))$, $t \in [0, T]$. Rewrite the energy conservation equation in the local coordinates $(U, \boldsymbol{\varXi})$ by using the kinetic matrix $g_{\ell,\ell'}$ in this reference system:

$$\frac{1}{2}\sum_{\ell',\ell''=1}^{Nd} g_{\ell',\ell''}(\mathbf{b}_\lambda(t))\dot{b}_{\lambda\,\ell'}(t)\dot{b}_{\lambda\,\ell''}(t) + V^{(a)}(\boldsymbol{\varXi}(\mathbf{b}_\lambda(t))) + \frac{\lambda}{2}\sum_{i=1}^{s} b_{\lambda\,i}(t)^2 = E, \tag{3.8.16}$$

having used Eq. (3.8.11) to express W.

The first of the above three addends is non-negative (being the kinetic energy; see, also, Proposition 11,§3.7, p.182). Hence, Eq. (3.8.16) implies

$$|b_{\lambda,j}(t)| \le \big(\frac{2C}{\lambda}\big)^{\frac{1}{2}}, \qquad j = 1, 2, \ldots, s. \tag{3.8.17}$$

if C is defined by Eq. (3.8.14).

From the examples discussed in §3.6, it is expected that the motion $t \to \mathbf{b}_\lambda(t)$, although squeezed on $\varSigma$, will very quickly oscillate transversally to $\varSigma$. It will therefore be useful to estimate the velocities $\dot{b}_{\lambda\,i}(t)$, $i = 1, \ldots, s$, of the "vanishing coordinates". Equation (3.8.16) also provides such estimates: by

Proposition 11, p.182, we can say that there is a constant $g^{-1} = \{$minimum of $C(\boldsymbol{\beta})$ in Eq. (3.7.29) for $\boldsymbol{\Xi}(\boldsymbol{\beta}) \in S_\varrho\}$ for which

$$\frac{1}{2} g^{-1} \sum_{\ell=1}^{Nd} (\dot{b}_{\lambda\, i}(t))^2 \le \frac{1}{2} \sum_{\ell',\ell''=1}^{Nd} g_{\ell',\ell''}(\mathbf{b}_\lambda(t)) \dot{b}_{\lambda\, \ell'}(t) \dot{b}_{\lambda\, \ell''}(t), \tag{3.8.18}$$

$\forall\, t \in [0,T]$, because for such t's and by the choice of T, $\mathbf{x}_\lambda(t) = \boldsymbol{\Xi}(\mathbf{b}_\lambda(t)) \in S_\varrho$.

Then Eqs. (3.8.18) and (3.8.16) imply, $\forall \lambda \ge 0$,

$$|\dot{b}_{\lambda,\ell}(t)| \le \sqrt{2Cg}, \qquad \ell = 1, \ldots, Nd,\; t \in [0,T]. \tag{3.8.19}$$

By the orthogonality and adaptation properties of the coordinate system $(U, \boldsymbol{\Xi})$, setting $\boldsymbol{\beta} = (\boldsymbol{\beta}_v, \boldsymbol{\beta}_n)$ with $\boldsymbol{\beta}_v \stackrel{def}{=} (\beta_1, \ldots, \beta_s) \in \mathcal{R}^s$, $\boldsymbol{\beta}_n = (\beta_{s+1}, \ldots, \beta_{Nd}) \in \mathcal{R}^{Nd-s}$, it is

$$g_{\ell',\ell''}(\mathbf{0}, \boldsymbol{\beta}_n) \equiv 0, \qquad \ell' = 1, \ldots, s; \ell'' = s+1, \ldots Nd \tag{3.8.20}$$

(orthogonality) and

$$g_{\ell,\ell'}(\mathbf{0}, \boldsymbol{\beta}_n) \equiv \gamma_{\ell,\ell'}, \qquad \ell, \ell' = 1, \ldots, s, \tag{3.8.21}$$

(good adaptation), where γ is a constant $s \times s$ matrix. Since the functions $g_{\ell',\ell''}(\boldsymbol{\beta})$, $\boldsymbol{\beta} \in \Omega$ are C^∞ functions, the Taylor-Lagrange theorem (see Appendix B), $\forall\, (\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) \in \Omega$, $\forall\, \ell = 1, \ldots, s$, $\ell' = s+1, \ldots, Nd$, yields

$$g_{\ell\,\ell'}(\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) = \sum_{j=1}^{s} g_{\ell\,\ell',j}(\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) \beta_j \tag{3.8.22}$$

and $\forall\, \ell, \ell' = 1, \ldots, s$:

$$g_{\ell\,\ell'}(\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) = \gamma_{\ell\,\ell'} + \sum_{j=1}^{s} g_{\ell\,\ell',j}(\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) \beta_j, \tag{3.8.23}$$

where $g_{\ell\,\ell',j}(\boldsymbol{\beta})$, $\boldsymbol{\beta} \in \Omega$, are suitable $C^\infty(\Omega)$ functions.

Equations (3.8.22) and (3.8.23) can be used to write “more explicitly” the equations of motion (3.7.15) for the “non constrained coordinates”, i.e., for the β_j's with $j = s+1, \ldots, Nd$. Using Eq. (3.8.1), one finds, for $\ell = s+1, \ldots . Nd$:

$$\begin{aligned} &\frac{d}{dt} \Big\{ \Big[\sum_{\ell'=1}^{s} \sum_{j=1}^{s} g_{\ell\,\ell',j}(b_\lambda(t)) b_{\lambda\, j}(t) \dot{b}_{\lambda\, \ell'}(t) \Big] + \sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(b_\lambda(t)) \dot{b}_{\lambda\, \ell'}(t) \Big\} \\ &= - \sum_{k=1}^{N} \frac{\partial V^{(a)}}{\partial \boldsymbol{\xi}^{(k)}} (\boldsymbol{\Xi}(\mathbf{b}_\lambda(t))) \cdot \frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \beta_\ell} (\mathbf{b}_\lambda(t)) \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\sum_{\ell',\ell''}^{s+1,Nd}\frac{\partial g_{\ell'\,\ell''}}{\partial\beta_\ell}(\mathbf{b}_\lambda(t))\dot b_{\lambda\,\ell'}(t)\dot b_{\lambda\,\ell''}(t)\\
&+\Bigg[\sum_{\ell'=1}^{s}\sum_{\ell''=s+1}^{Nd}\sum_{j=1}^{s}\frac{\partial g_{\ell,\ell',j}}{\partial\beta_\ell}(\mathbf{b}_\lambda(t))b_{\lambda\,j}(t)\dot b_{\lambda\,\ell'}(t)\dot b_{\lambda\,\ell''}(t) \qquad (3.8.24)\\
&+\frac{1}{2}\sum_{\ell',\ell''=1}^{s}\sum_{j=1}^{s}\frac{\partial g_{\ell,\ell',j}}{\partial\beta_\ell}(\mathbf{b}_\lambda(t))b_{\lambda\,j}(t)\dot b_{\lambda\,\ell'}(t)\dot b_{\lambda\,\ell''}(t)\Bigg]
\end{aligned}
$$

where the square brackets isolate the terms which should vanish, as $\lambda \to +\infty$, in order that Eq. (3.8.24) could reduce at least formally to Eq. (3.7.20) as wished on the basis of Definition 13,§3.7, p.180 and Observation (4) to Definition 13 p.182.

Note that in Eq. (3.8.24) every term in square brackets contains factors proportional to one of the first s coordinates which, by Eq. (3.8.17), vanish as $\lambda \to +\infty$ uniformly in $t \in [0, T]$. The coefficients in Eq. (3.8.24) of such coordinates are uniformly bounded in λ, as the motion takes place in S_ϱ, for $t \in [0, T]$, and there the $g\ldots$ are C^∞ functions and, therefore, bounded together with their derivatives; furthermore, Eq. (3.8.19) provides λ-independent bounds for $\dot b_{\lambda\,\ell}(t)$.

To understand in a rigorous way that the above formal convergence of Eqs. (3.8.17) and (3.8.24) to Eqs. (3.7.8) and (3.7.20) implies that, uniformly in $t \in [0, T]$, the functions $t \to b_{\lambda\,\ell}(t)$, $\ell = s+1, \ldots, Nd$, converge to limits $b_\ell(t)$ verifying Eq. (3.7.20) with the desired initial conditions (3.7.24), some more work is still necessary.

Integrate both sides of Eq. (3.8.20) with respect to $t, \forall\, \ell = s+1, \ldots, Nd$:

$$
\begin{aligned}
&\sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(\mathbf{b}_\lambda(t))\dot{b}_{\lambda\,\ell'}(t) + \Big[\sum_{\ell',j=1}^{s} g_{\ell,\ell',j}(\mathbf{b}_\lambda(t))\, b_{\lambda\,j}(t)\dot{b}_{\lambda\,\ell'}(t)\Big] \\
&- \sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(\mathbf{b}(0))\,\dot{b}_{\ell'}(0) \\
&= \int_0^t \Big\{ -\sum_{k=1}^{N} \frac{\partial V^{(a)}}{\partial \xi^{(k)}}(\boldsymbol{\Xi}(\mathbf{b}_\lambda(t'))) \cdot \frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \beta_\ell}(\mathbf{b}_\lambda(t')) \\
&+ \frac{1}{2}\sum_{\ell',\ell''=s+1}^{Nd} \frac{\partial g_{\ell'\,\ell''}}{\partial \beta_\ell}(\mathbf{b}_\lambda(t'))\dot{b}_{\lambda\,\ell'}(t')\dot{b}_{\lambda\,\ell''}(t')\Big\}\, dt' \\
&+ \Big[\int_0^t dt' \Big\{ \sum_{j=1}^{s}\sum_{\ell''=s+1}^{Nd}\sum_{j=1}^{Nd} \frac{\partial g_{\ell',\ell'',j}}{\partial \beta_\ell}(\mathbf{b}_\lambda(t'))b_{\lambda\,\ell}(t')\dot{b}_{\lambda\,\ell'}(t')\dot{b}_{\lambda\,\ell''}(t') \\
&+ \frac{1}{2}\sum_{\ell',\ell''=1}^{s}\sum_{j=1}^{s} \frac{\partial g_{\ell',\ell'',j}}{\partial \beta_\ell}(\mathbf{b}_\lambda(t'))b_{\lambda\,\ell}(t')\dot{b}_{\lambda\,\ell'}(t')\dot{b}_{\lambda\,\ell''}(t')\Big],
\end{aligned}
\tag{3.8.25}
$$

where in the second line the hypothesis that $b_{\lambda\,\ell}(0) = 0$, if $\ell = 1, \ldots, s$, is used together with the λ-independence of the initial data $b_{\lambda\,\ell}, \dot{b}_{\lambda\,\ell}$, $\ell = 1, \ldots, Nd$.

Now bring the second and third addends to the right-hand side and consider the resulting equations as $(Nd - s)$ linear equations in the $(Nd - s)$ unknowns $b_{\lambda\,\ell}(t)$, $\ell = s + l, \ldots, Nd$, pretending that the right-hand side is known. The matrix of the coefficients is the last $(Nd - s) \times (Nd - s)$ principal submatrix g_s, of the kinetic matrix g: $(g_s)ij = g_{ij}(\mathbf{b}_\lambda(t))$, $i, j = s+1, \ldots, Nd$. By Proposition 11,§3.7, p.182, g_s admits an inverse matrix $d_s(\mathbf{b}_\lambda(t))^{-1}$ (making explicit its $\mathbf{b}_\lambda(t)$ dependence). Therefore, $\dot{b}_{\lambda\,\ell}(t)$, $\ell = s + 1, \ldots, Nd$, can be expressed in terms of the right-hand side. Thus:

$$
\dot{b}_{\lambda\,\overline{\ell}}(t) = \Big[-\sum_{\ell=s+1}^{Nd} (g_s^{-1}(\mathbf{b}_\lambda(t)))_{\overline{\ell}\,\ell} \sum_{\ell'=1}^{s}\sum_{j=1}^{s} g_{\ell',\ell,j}(\mathbf{b}_\lambda(t))\, b_{\lambda\,j}(t)\dot{b}_{\lambda\,\ell}(t)\Big]
$$

$$
\begin{aligned}
&+\sum_{\ell=s+1}^{Nd}(g_s^{-1}(\mathbf{b}_\lambda(t)))_{\overline{\ell}\,\ell}\sum_{\ell'=s+1}^{Nd}g_{\ell\,\ell'}(\mathbf{b}(0))\dot{\mathbf{b}}(0)+\\
&+\sum_{\ell=s+1}^{Nd}(g_s^{-1}(\mathbf{b}_\lambda(t)))_{\overline{\ell}\,\ell}\\
&\cdot\int_0^t\Big\{\Big(-\sum_{k=1}^{N}\frac{\partial V^{(a)}(\boldsymbol{\Xi}(\mathbf{b}_\lambda(t')))}{\partial\boldsymbol{\xi}^{(k)}}\cdot\frac{\partial\boldsymbol{\Xi}^{(k)}(\mathbf{b}_\lambda(t'))}{\partial\beta_\ell}\\
&+\frac{1}{2}\sum_{\ell',\ell''=s+1}^{Nd}\frac{\partial g_{\ell'\,\ell''}(\mathbf{b}_\lambda(t'))}{\partial\beta_\ell}\dot{b}_{\lambda\,\ell'}(t)\dot{b}_{\lambda\,\ell''}(t)\Big)\Big\}\,dt'\\
&+\Big[\sum_{\ell=s+1}^{Nd}(g_s^{-1}(\mathbf{b}_\lambda(t)))_{\overline{\ell}\,\ell}\\
&\cdot\int_0^t\Big\{\sum_{\ell'=1}^{s}\sum_{\ell''=s+1}^{Nd}\sum_{j=1}^{s}\frac{\partial g_{\ell',\ell'',j}(\mathbf{b}_\lambda(t'))}{\partial\beta_\ell}\,b_{\lambda\,j}(t')\dot{b}_{\lambda\,\ell'}(t')\dot{b}_{\lambda\,\ell''}(t')\\
&+\frac{1}{2}\sum_{\ell',\ell''=1}^{s}\sum_{j=1}^{s}\frac{\partial g_{\ell',\ell'',j}(\mathbf{b}_\lambda(t'))}{\partial\beta_\ell}\,b_{\lambda\,j}(t')\dot{b}_{\lambda\,\ell'}(t')\dot{b}_{\lambda\,\ell''}(t')\Big\}\,dt'\Big]
\end{aligned}
\tag{3.8.26}
$$

It has, now, to be remarked that the terms in square brackets vanish uniformly in $t \in [0,T]$ as $\lambda \to +\infty$ because of Eqs. (3.8.17) and (3.8.19) and because of the uniform boundedness in $\boldsymbol{\Xi}^{-1}(S_\varrho)$ of the g functions and of their derivatives.

Furthermore, convergence to a limit, as $\lambda \to +\infty$ of the terms which are not in square brackets in Eq. (3.8.26) follows: call $\delta_{\lambda,\overline{\ell}}(t)$, $t \in [0,T]$ their sum and show, first, that a subsequence $\lambda_n \to +\infty$, extracted from an arbitrary diverging sequence, exists such that $\delta_{\lambda_n,\overline{\ell}}(t)$ converges to a limit $\delta_{\overline{\ell}}(t)$ uniformly in $t \in [0,T], \forall\,\overline{\ell} = s+1,\ldots,Nd$.

This will be shown by proving that the family of functions on $[0,T]$ parameterized by λ and $\overline{\ell}$: $(\delta_{\lambda,\ell})_{\lambda\ge 1,\,\ell=s+1,\ldots,Nd}$ is an equicontinuous and equibounded family of functions on $[0,T]$, and then applying the Ascoli-Arzelà theorem (see Appendix H).

Finally, the actual existence of the limit as $\lambda \to +\infty$ of $\delta_{\lambda,\ell}(t)$ will be obtained by showing that every limit of the converging subsequences verifies a certain differential equation with given initial conditions, whatever the subsequence is, and applying the uniqueness theorem for differential equations: the equation will essentially turn out to coincide with Eq. (3.7.20) and the proof will then be complete.

Equiboundedness (see Appendix H) of the functions is clear from Eqs. (3.8.17) and (3.8.19). Equicontinuity of the contribution to $\delta_{\lambda,\overline{\ell}}$ coming from the integral of Eq. (3.8.26) and that coming from the part outside the integral can be separately shown. They follow from the remarks:

(i) Consider a family $(\mu_\alpha)_{\alpha\in A}$ of functions on $[0,T]$ given by

$$\mu_\alpha(t) = \int_0^t \nu_\alpha(\tau)\, d\tau \tag{3.8.27}$$

where $(\nu_\alpha)_{\alpha\in A}$ is a family of equibounded continuous functions on $[0,T]$, i.e., a family of functions bounded as $|\nu_\alpha(t)| \le B$, $\forall\, t \in [0,T]$, with a suitable B, $\forall\, \alpha \in A$. Then the family $(\mu_\alpha)_{\alpha\in A}$ is equicontinuous; in fact,

$$|\mu_\alpha(t) - \mu_\alpha(t')| = \Big| \int_{t'}^t \nu_\alpha(\tau)\, d\tau \Big| \le\ B\, |t-t'|. \tag{3.8.28}$$

(ii) Families of functions obtained by composing a given $C^\infty(\mathcal{R}^h)$ function and a family of equicontinuous equibounded $\mathcal{R}^h$-valued functions on $[0,T]$ form equicontinuous equibounded families of functions $\forall h > 0$ (exercise).

By suitably combining the criteria (i) and (ii), Eqs. (3.8.17), and (3.8.19), the fact that $(g^{-1})_{\ell'\,\ell''}$ are C^∞ functions on $\boldsymbol{\Xi}^{-1}(S_\varrho)$ and $t \to \mathbf{b}_\lambda(t)$ is an equicontinuous family [by (i) and by Eq. (3.8.19)] one realizes that $\delta_{\lambda\,\overline{\ell}}$ form an equicontinuous equibounded family of functions on $[0,T]$ parameterized by $\lambda \ge 1, \overline{\ell} = s+1,\ldots,Nd$.

Then the Ascoli-Arzelà criterion (see Appendix H) states that from every diverging sequence of positive numbers, it is possible to extract a diverging subsequence $(\lambda_n)_{n\in\mathcal{Z}_+}$ such that the limit

$$\lim_{n\to\infty} \delta_{\lambda_n\,\overline{\ell}}(t) = \delta_{\overline{\ell}}(t) \tag{3.8.29}$$

exists uniformly for $t \in [0,T], l = s+1,\ldots,Nd$.

Equation (3.8.26) then implies (since it has already been observed that the terms in square brackets in the right-hand side vanish uniformly as $\lambda \to +\infty$, $\forall\, t \in [0,T]$) that

$$\lim_{n\to\infty} \dot{\delta}_{\lambda_n\,\overline{\ell}}(t) = \dot{\delta}_{\overline{\ell}}(t). \qquad \text{Hence,} \tag{3.8.30}$$

$$\begin{aligned} \lim_{n\to\infty} b_{\lambda_n,\overline{\ell}}(t) &\equiv \lim_{n\to\infty} \Big(b_{\lambda_n,\overline{\ell}}(0) + \int_0^t \dot{b}_{\lambda_n\,\overline{\ell}}(\tau)\, dt \Big) \\ &= b_{\overline{\ell}}(0) + \int_0^t \dot{\delta}_{\overline{\ell}}(\tau)\, d\tau \stackrel{def}{=} b_{\overline{\ell}}(t), \qquad \overline{\ell} = s+1,\ldots,Nd, \end{aligned} \tag{3.8.31}$$

uniformly in $t \in [0,T]$, because the initial datum is λ independent, and $b_{\overline{\ell}}(t)$ is defined by the last identity. Of course, by changing the subsequence λ_n we cannot yet be sure that $\delta_{\overline{\ell}}$ and $b_{\overline{\ell}}$, thus defined, do not change.

The functions $t \to b_{\overline{\ell}}(t)$, $t \in [0,T]$, defined in Eq. (3.8.31) are, by Eq. (3.8.31) itself, once differentiable and

$$\dot{b}_{\overline{\ell}}(t) = \delta_{\overline{\ell}}(t), \qquad \forall\, t \in [0,T], \quad \forall\, i = s+1,\ldots,Nd. \tag{3.8.32}$$

Coming back to Eq. (3.8.25) with $\lambda = \lambda_n$ and using Eqs. (3.8.17), (3.8.31), (3.8.30), and (3.8.32), we find that as $n \to \infty$, $\forall\, \ell = s+1,\ldots,Nd$:

$$\begin{aligned}\sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(\mathbf{b}(t))\dot{b}_{\ell'}(t) &= \sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(\mathbf{b}(0))\dot{b}_{\ell'}(0)\\ &+ \int_0^t \Big\{ -\sum_{k=1}^{N} \frac{\partial V^{(a)}}{\partial \xi^{(k)}}(\boldsymbol{\Xi}(\mathbf{b}(t'))) \cdot \frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \beta_\ell}(\mathbf{b}(t'))\\ &+ \frac{1}{2}\sum_{\ell',\ell''=s+1}^{Nd} \frac{\partial g_{\ell'\,\ell''}}{\partial \beta_\ell}(\mathbf{b}(t'))\dot{b}_{\ell'}(t')\dot{b}_{\ell'}(t'')\Big\}\, dt',\end{aligned} \tag{3.8.33}$$

having set $\mathbf{b}(t) = (0, \ldots, 0, b_{s+1}(t), \ldots, b_{Nd}(t))$ [recall that $b_\ell(t)$ is defined by Eq. (3.8.31) only for $\ell = s+1, \ldots, Nd$]. Therefore $t \to \mathbf{b}(t)$ verifies the wanted initial conditions at $t = 0$, Eq. (3.7.24), as well as Eq. (3.7.18) and, by differentiating Eq. (3.8.33), also Eq. (3.7.20). It is also true that $\mathbf{b} \in C^\infty([0,T])$. In fact, pretending that the right-hand side of Eq. (3.8.33) is known, we interpret Eq. (3.8.33) as a linear system in the unknowns $\dot{b}_\ell(t)$: its coefficients form the already-met nonsingular matrix $g_s(\mathbf{b}(t))$. Proceeding as in Eq. (3.8.26), we find

$$\begin{aligned}\dot{b}_{\overline{\ell}}(t) &= \sum_{\ell=s+1}^{Nd} (g_s^{-1}(\mathbf{b}(t)))_{\overline{\ell}\,\ell}\Big(\sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(\mathbf{b}(0))\dot{b}_{\ell'}(0)\\ &+ \int_0^t \Big\{ -\sum_{k=1}^{N} \frac{\partial V^{(a)}}{\partial \xi^{(k)}}(\boldsymbol{\Xi}(\mathbf{b}(t'))) \cdot \frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \beta_\ell}(\mathbf{b}(t'))\\ &+ \frac{1}{2}\sum_{\ell',\ell''=s+1}^{Nd} \frac{\partial g_{\ell'\,\ell''}}{\partial \beta_\ell}(\mathbf{b}(t'))\dot{b}_{\ell'}(t')\dot{b}_{\ell'}(t'')\Big\}\, dt'\Big),\end{aligned} \tag{3.8.34}$$

and since the right-hand side is obviously once differentiable, it follows that b_ℓ; is twice differentiable. Differentiating both sides, we find an expression for $\ddot{b}_\ell$ in terms of $\mathbf{b}, \dot{\mathbf{b}}$, and some integrals: hence, $\dot{\mathbf{b}}$ is differentiable, etc. So $\mathbf{b}$ is in $C^\infty([0,T])$.

It remains to show that the limit as $\lambda \to +\infty$ of $b_{\lambda\,\overline{\ell}}(t)$ exists, $\forall\,\overline{\ell} = s+1, \ldots, Nd$, without "passing to subsequences". It suffices to show that every divergent subsequence $\lambda_n \to +\infty$ for which the limit $\lim_{n\to+\infty} b_{\lambda_n\,\overline{\ell}}(t)$, $\overline{\ell} = s+1, \ldots, Nd$, exists uniformly has to converge to the same limit.

It is enough to show that there is only one function $t \to \mathbf{b}(t)$ verifying Eq. (3.8.33) and in $C^\infty([0,T])$ and such that $\mathbf{b}(t) \in \boldsymbol{\Xi}^{-1}(S_\varrho)$, $\forall\, t \in [0,T]$, because every limit of a uniformly convergent subsequence has to verify Eq. (3.8.33). The following trick, which will be sublimated in §3.11 and§3.12, can be used.

Set, $\forall\,\ell = s+1, \ldots .Nd$:

$$p_\ell(t) \stackrel{def}{=} \sum_{\ell'=s+1}^{Nd} g_{\ell\,\ell'}(\mathbf{b}(t))\dot{b}_{\ell'}(t) \tag{3.8.35}$$

or, for short,

$$\mathbf{p} = g_s(\mathbf{b})\dot{\mathbf{b}}, \tag{3.8.36}$$

where $g_s(\boldsymbol{\beta})$ is the last principal matrix of $g(\boldsymbol{\beta})$ of order $Nd - s$. Then, by differentiation with respect to t, Eq. (3.8.33) yields the following equations, $\forall\, \ell = s+1, \ldots, Nd$:

$$\begin{aligned} \dot{p}_\ell =& -\sum_{k=1}^{N} \frac{\partial V^{(a)}}{\partial \xi^{(k)}}(\boldsymbol{\Xi}(\mathbf{b})) \cdot \frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \beta_\ell}(\mathbf{b}) \\ &+ \frac{1}{2} \sum_{\ell',\ell''=s+1}^{Nd} \frac{\partial g_{\ell'\,\ell''}}{\partial \beta'_\ell}(\mathbf{b})\,(g_s(\mathbf{b})^{-1}\mathbf{p})_{\ell''}\,(g_s(\mathbf{b})^{-1}\mathbf{p})_{\ell''}. \\ \dot{b}_\ell =& (g_s(\mathbf{b})^{-1}\mathbf{p})_\ell \end{aligned} \tag{3.8.37}$$

with the notations of Eq. (3.8.33), having dropped the t dependence from $\mathbf{p}(t), \mathbf{b}(t)$ and having deduced the second equation from Eq. (3.8.36). In Eq. (3.8.37), $\mathbf{b}$ means $(0, \ldots, 0, b_{s+1}, \ldots, b_{Nd})$.

Note that $p_\ell(0), b_\ell(0)$, $\ell = s+1, \ldots, Nd$, are, by Eq. (3.8.36) or by assumption, independent of the sequence used to construct $\mathbf{b}$.

So every sequence $\lambda_n \to +\infty$, for which $b_{\lambda_n\,\ell}(t)$ is uniformly convergent in $t \in [0,T]$ to a limit, $\forall\, \ell = s+1, \ldots, Nd$, can be used to construct a solution of the differential equation (3.8.37) for $t \to (p_\ell(t), b_\ell(t))_{\ell=s+1,\ldots,Nd}$ verifying the initial condition $(p_\ell(0), b_\ell(0))_{\ell=s+1,\ldots,Nd}$.

Eq. (3.8.37) is not quite a differential equation of the type considered in the uniqueness theorem, Proposition 1, §2.2, p.14, since the right-hand side of Eq. (3.8.37) is defined only for $\mathbf{b} \in \boldsymbol{\Xi}^{-1}(S_\varrho)$ as a function of the p_ℓ,'s, b_ℓ's. However, all functions $\mathbf{b}(t)$, $t \in [0,T]$, which can be built via the above construction, are such that $\mathbf{b}(t) \in \boldsymbol{\Xi}^{-1}(S_\varrho)$, $t \in [0,T]$. Then easy from Proposition 1, p.14, as a corollary, it follows that every solution to Eq. (3.8.37) $t \to (\mathbf{p}(t), \mathbf{b}(t))$, $t \in [0,T]$, verifying $\mathbf{b}(t) \in \boldsymbol{\Xi}^{-1}(S_\varrho), \forall\, t \in [0,T]$, must be identical to every other with this property. mbe

3.8.1 Problems

1. Let $\Sigma \subset \mathcal{R}^{Nd}$ be a regular surface with codimension s. Let $(U, \boldsymbol{\Xi})$ be a regular system of local coordinates well adapted and orthogonal on Σ: with respect to the scalar product of Eq. (3.7.1). Denote $\boldsymbol{\beta} \in \Omega$ the coordinates of $\boldsymbol{\xi} = \boldsymbol{\Xi}(\boldsymbol{\beta}) \in U$. Set $\boldsymbol{\beta} = (\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) \in \mathcal{R}^s \times \mathcal{R}^{Nd-s}$. Show that the change of coordinates $(\boldsymbol{\beta}_v, \boldsymbol{\beta}_n) \to (\Lambda\boldsymbol{\beta}_v, \widetilde{\Lambda}\boldsymbol{\beta}_n)$, with Λ and $\widetilde{\Lambda}$ being two $s \times s$ and $(Nd-s) \times (Nd-s)$ constant matrices, allows us to define a new system of local coordinates which is still well adapted and orthogonal on Σ.

2. In the context of Problem 1, let $\overline{W}(\beta_1, \ldots, \beta_s)$ be a C^∞ function of $(\beta_1, \ldots, \beta_s, \beta_{s+1}, \ldots, \beta_{sNd})$ independent of the last $(Nd-s)$ coordinates. Suppose that $M_{ij} = \frac{\partial^2 \overline{W}}{\partial \beta_i \partial \beta_j}(\mathbf{0})$, $i, j = 1, ..., s$, is a $s \times s$ matrix which is positive definite. Show that there is a change of coordinates $\boldsymbol{\beta} \to \boldsymbol{\beta}'$ of the linear type considered in Problem 1 changing $\overline{W}$ into a function

such that $M_{ij} = \delta_{ij}$, $i, j = 1, .., s$. (*Hint:* Use $\widetilde{\Lambda} = 1, \Lambda = J = \{$orthogonal matrix diagonalizing $M\}$ (see Appendix F, F.4). Then make a further change of coordinates of the same type with $\widetilde{\Lambda} = 1$ and $\Lambda = \{$diagonal matrix with diagonal elements $(w_1^{-\frac{1}{2}}, \ldots, w_s^{-\frac{1}{2}})$ where $(w_1, \ldots, w_s)$ are the s eigenvalues of M.)

3. Show that there is essentially no change in the proof of Proposition 13 if Eq. (3.8.11) is changed into

$$\overline{W}(\beta_1, \ldots, \beta_s) = \frac{1}{2}\sum_{i=1}^{s} \beta_i^2 + o(\beta_1^2 + \ldots + \beta_s^2).$$

4.* Alternatively to Problem 3, but with the same assumptions, show that there is a change of coordinates which, possibly restricting the size of U to $U' \subset U$, changes $\boldsymbol{\beta}$ into $\boldsymbol{\beta}'$, retaining the orthogonality and good adaptation of the $\boldsymbol{\beta}'$ coordinates, changing $\widetilde{W}$ into $\frac{1}{2}\sum_{j=1}^{s} \beta_j'^2$. (*Hint:* For $\boldsymbol{\beta} \in \Omega$, $\Omega \equiv$ (basis of $(U\boldsymbol{\Xi})$), let $W(\boldsymbol{\beta}) = \frac{1}{2}\boldsymbol{\beta}^2 + \sum_{i,j,\ell}^{1,s} \gamma_{ij\ell}(\boldsymbol{\beta})\beta_i\beta_j\beta_\ell$, where $\gamma_{ij\ell}$ are suitable ${}^\infty(\Omega)$ functions symmetric in the indices i, j, ℓ (this assumption is not restrictive because of Problem 2 and of the Taylor-Lagrange theorem, Appendix B). Then define $\beta_\ell' = \beta_\ell + \sum_{j,k=1}^{s} f_{\ell,j,k}(\boldsymbol{\beta})\beta_j\beta_k$ with f symmetric in i, j, k and of class C^∞ in $\boldsymbol{\beta}$ and impose

$$\begin{aligned}
&\frac{1}{2}\boldsymbol{\beta}^2 + \sum_{i,j,\ell}^{1,s} \gamma_{i,j,\ell}(\boldsymbol{\beta})\beta_i\beta_j\beta_\ell \equiv \frac{1}{2}\boldsymbol{\beta}'^2 \\
&= \frac{1}{2}\boldsymbol{\beta}^2 + \sum_{\ell,j,k=1}^{s} \beta_\ell f_{\ell,j,k}(\boldsymbol{\beta})\beta_j\beta_k + \frac{1}{2}\sum_{\ell,j,k,j',k'} f_{\ell,j,k}(\boldsymbol{\beta}) f_{\ell,j',k'}(\boldsymbol{\beta})\beta_\ell^2\beta_j\beta_k\beta_{j'}\beta_{k'}
\end{aligned}$$

Therefore, $\gamma_{jk\ell}$ has to be equal to

$$f_{j,k,\ell}(\boldsymbol{\beta}) + \frac{1}{2}\sum_{(\ell_1,j_1,k_1,j_1',k_1')\supset^*(j,k,\ell)} f_{\ell_1,j_1,k_1}(\boldsymbol{\beta}) f_{\ell_1,j_1',k_1'}(\boldsymbol{\beta}) \left[\frac{\beta_{\ell_1}^2\beta_{j_1}\beta_{k_1}\beta_{j_1'}\beta_{k_1'}}{\beta_j\beta_k\beta_\ell}\right],$$

where $\supset^*$ means that the monomial in square brackets has to "simplify" so that all the terms in the denominator cancel with some in the numerator. Show that for small $\boldsymbol{\beta}$, by the implicit functions theorem, the above relation allows us to determine f in terms of γ and $\boldsymbol{\beta}$. Then, again by the implicit functions theorem, invert the relation between $\boldsymbol{\beta}$ and $\boldsymbol{\beta}'$ to complete the change of coordinates.)

5.* Let $t \to \mathbf{x}(t)$, $t \in \mathcal{R}_+$, be a motion of N points, with masses $m_1, \ldots, m_N > 0$, which develops under the influence of an active force $\mathbf{F}^{(a)}$, conservative with potential energy $V^{(a)} \in C^\infty(\mathcal{R}^{Nd})$ bounded from below, and of an ideal constraint to a regular surface $\Sigma \subset \mathcal{R}^{Nd}$ with a codimension s.

Let $\boldsymbol{\xi}_0 = \mathbf{x}(0)$, $\boldsymbol{\eta}_0 = \dot{\mathbf{x}}(0)$, and let $\widetilde{\boldsymbol{\eta}}_0$ be any velocity vector such that $\widetilde{\boldsymbol{\eta}}_0^{\Sigma} = \boldsymbol{\eta}_0$ and call $t \to \mathbf{x}_\lambda(t)$ the motion with initial datum $(\widetilde{\boldsymbol{\eta}}_0, \boldsymbol{\xi}_0)$ of the same system moving under the influence of the same active force and of an approximate constraint model (Σ, W, λ) verifying the assumptions of Proposition 13. Call $T_0 = \{$supremum of the T's such that $\lim_{\lambda\to+\infty} \mathbf{x}_\lambda(t) = \mathbf{x}(t)$ uniformly for $t \in [0, T]\}$. Show that $T_0 = +\infty$. (*Hint:* The part of Proposition 13 proved in this section says that $T_0 > 0$. Suppose $T_0 < +\infty$. Let $\widehat{\boldsymbol{\xi}}_0 = \mathbf{x}(T_0)$, $\widehat{\boldsymbol{\eta}}_0 = \dot{\mathbf{x}}(T_0)$ and note that the energies of the motions $\mathbf{x}_\lambda$ and $\mathbf{x}$ are λ independent and coincide. Discuss the system's motion in the coordinate system $(U, \boldsymbol{\Xi})$ around $\widehat{\boldsymbol{\xi}}_0$ which is well adapted and orthogonal to Σ and in which W admits a representation like Eq. (3.8.11). Show that in $(U, \boldsymbol{\Xi})$, $\widehat{\mathbf{x}}_\lambda$ verifies equations like Eqs. (3.8.24), (3.8.25), (3.8.26), and (3.8.31)

with some slight changes which do not affect the conclusion that $\lim_{\lambda\to+\infty} \mathbf{x}_\lambda(t) = \mathbf{x}(t)$ as long as $\mathbf{x}_\lambda(t)$ stays inside U at a positive distance from ∂U. Since the conservation of energy implies the bound of Eq. (3.8.14) on speed, it is clear that for large λ, $\mathbf{x}_\lambda(t)$ and $\mathbf{x}(t)$ will stay at a positive distance from ∂U for all times in a neighborhood of T_0. Hence, $T_0 = +\infty$.)

6.* Show that a system of N point masses, with masses $m_1, ..., m_N > 0$, bound by an ideal bilateral constraint to a regular surface $\Sigma \subset \mathcal{R}^{Nd}$ and also subject to a conservative active force with inferiorly bounded potential energy $V^{(a)}$, has (for all $\boldsymbol{\xi}_0 \in \Sigma$ and $\boldsymbol{\eta}_0$ tangent to Σ) a unique global motion $t \to \mathbf{x}(t)$, $t \in \mathcal{R}_+$, such that $\dot{\mathbf{x}}(0) = \boldsymbol{\eta}_0, \mathbf{x}(0) = \boldsymbol{\xi}_0$; i.e., a motion verifying Eq. (3.7.20) in every local system of regular coordinates (*Hint*: Use the energy conservation and the existence and uniqueness theorems for Eq. (3.7.20) following from its transformation into Eq. (3.8.37): energy conservation together with the semiboundedness of $V^{(a)}$ gives an a priori estimate.)

3.9 Application to Rigid Motion. König's Theorem

The general perfection criterion for approximate constraints discussed in §3.8 is interesting because it establishes perfection of some classes of constraint models.

In this section, as an application of the results of §3.8, Proposition 13, it will be shown that a natural rigidity constraint model is approximately ideal.

Consider the following model (Σ, W, λ), which is one of the most important constraint models for N points. Let $\ell_{ij} > 0$ be given numbers defined for $(i, j) \in S = \{$subset of the set of pairs of different points in $(1, \dots, N)\}$; let σ_i, $i \in T \subset \{1, \dots, N\}$, be a family of regular surfaces in $\mathcal{R}^3$; then define

$$\begin{aligned}\Sigma = \{\boldsymbol{\xi} \mid \boldsymbol{\xi} = (\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(N)}) \in \mathcal{R}^{3N},\ &|\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}| = \ell_{ij} \text{ for } i, j \in S;\\ \boldsymbol{\xi}^{(i)} \in \sigma_i,\ \text{ for } i \in T\},&\end{aligned} \tag{3.9.1}$$

$$W(\boldsymbol{\xi}) = \sum_{i,j\in S} \psi_{ij}(|\boldsymbol{\xi}^{(i)} - \boldsymbol{\xi}^{(j)}| = \ell_{ij} + \sum_{i\in T} \psi_i(|\boldsymbol{\xi}^{(i)} - \sigma_i|^2), \tag{3.9.2}$$

where $\psi_{ij}, \psi_i \in C^\infty(\mathcal{R})$, $\psi_{ij}(0) = \psi_i(0) = 0$, and have a strict minimum at zero; the notation $|\boldsymbol{\xi} - \sigma_i|^2$ denotes a C^∞ function on $\mathcal{R}^3$ positive outside σ_i and near σ_i, equal to the square of the distance between $\boldsymbol{\xi}$ and σ_i. Here σ_i may also be a single point.

(Σ, W, λ) is a natural model of rigidity for some system points (those in S) and for permanence on a surface or on a point (if σ_i is zero dimensional) for some of the system points (those in T).

In applications, it is quite common to meet only constraints for which the above is a good model, when friction is neglected. It is not completely trivial to show that Eqs. (3.9.1) and (3.9.2) are an approximate ideal model in the sense of Definition 13, §3.7.

In this case, we shall examine, for simplicity, only the case in which Eq. (3.9.1)) is a "total rigidity" constraint, i.e., the case when S contains so many pairs (e.g. all) to allow only configurations which can be obtained by rigid motions of a single one or, at most, of finitely many. Nevertheless, we formulate the general result.

14 Proposition. *The model (Σ, W, λ) defined by Eqs. (3.9.1) and (3.9.2) is a model of an approximate ideal constraint for a system of N points with (arbitrary) masses $m_1, \ldots, m_N > 0$.*

Proof. (Case $T = \emptyset$, S such that Σ is a total rigidity constraint). The surface Σ, in the case under examination, decomposes into a finite number of connected parts, each representing a rigid system in the usual sense of the word.[9]

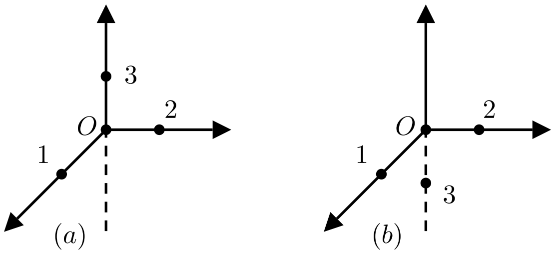


Fig.3.2

Fig.3.2. Example of two rigid disconnected configurations.

Suppose $N \geq 3$, the $N = 2$ case having been already discussed in the Example in §3.8, p.187. Suppose also that the points $1, 2$, and 3 are not aligned in the configurations of Σ: the degenerate case of N aligned points could be treated likewise.

The configurations $\boldsymbol{\xi}' \in \Sigma$ located on the same connected component of Σ shall be uniquely determined by the position G of the system baricenter in the "fixed" Cartesian reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ and by three orthogonal unit vectors $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ fixed with the system ("co-moving") and finally by the positions $P_1, \ldots, P_N$ of N points in the reference frame $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. By the rigidity constraint, the points $P_1, \ldots, P_N$ will have coordinates $(P_i - G)_\ell, \ell = 1, 2,\ i = 1, 2, \ldots, N$, which are given constants, $\forall\, \boldsymbol{\xi}'$ in the same connected component of Σ, in the frame $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. Suppose to have fixed $\mathbf{i}_3$ parallel to $(P_2 - P_1)$ and $\mathbf{i}_2$ parallel to the plane (P_1, P_2, P_3) (but orthogonal to $\mathbf{i}_3$).

To prove Proposition 14, it will be sufficient to build a system of coordinates, local near $\boldsymbol{\xi}_0 \in \Sigma$, regular, well adapted, and orthogonal to Σ with respect to the scalar product of Eq. (3.7.1) and with the extra property that W, the constraint structure function, has the property of Eq. (3.8.1).

[9] Σ my consist of several connected parts: for instance, if $N = 4$ and the distances of the points $0, 1, 2, 3$ are $d(0,1) = 1, d(0,2) = 1,\ d(1,2) = \sqrt{2}, d(1,3) = \sqrt{2},\ d(2,3) = \sqrt{2}$, respectively, then Σ contains two connected parts. The first consists of the configurations obtained by rotations and translations of the configuration in Fig.3.2(a) and the other of those in Fig. 3.2(b).

Without loss of generality, suppose that the plane of the first two axes in the co-moving frame $(G_0; \mathbf{i}_1^{(0)}, \mathbf{i}_2^{(0)}, \mathbf{i}_3^{(0)})$, associated with $\boldsymbol{\xi}_0 \in \Sigma$, i.e. the plane $(\mathbf{i}_1^{(0)}, \mathbf{i}_2^{(0)})$, is not parallel to the plane $(\mathbf{i}, \mathbf{j})$.

Let $\boldsymbol{\xi}$ be a configuration close to Σ: in general, $\boldsymbol{\xi} \notin \Sigma$ and with $\boldsymbol{\xi}$ are associated $3N$ coordinates, obtained through the following construction. $\boldsymbol{\xi}$ can be determined by assigning a configuration $\boldsymbol{\xi}' \in \Sigma$ and the vectors $\boldsymbol{\kappa}^{(1)}, \ldots, \boldsymbol{\kappa}^{(N)}$ providing the deviations of the points in $\boldsymbol{\xi}$ with respect to the corresponding points $P_1, \ldots, P_N$ in $\boldsymbol{\xi}'$. The $(3N+6)$ coordinates necessary to determine the $3N$ components of $\boldsymbol{\kappa}^{(1)}, \ldots, \boldsymbol{\kappa}^{(N)}$ in the frame $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ fixed with $\boldsymbol{\xi}'$ and the six coordinates giving the position and orientation in space of $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, i.e., of $\boldsymbol{\xi}'$, are redundant and six of them must be eliminated.

Coordinates that can be used to determine $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ are the three Cartesian coordinates of G in the fixed frame $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$ and the three "Euler angles" (θ, φ, ψ) of $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ where $\mathbf{n}$ is the unit vector along the intersection between the plane $(\mathbf{i}, \mathbf{j})$ and the plane $(\mathbf{i}_1, \mathbf{i}_2)$, arbitrarily oriented (the "node lines"). The angles $\theta = \widehat{\mathbf{i}_3\mathbf{k}}$, $\varphi = \widehat{\mathbf{i}\,\mathbf{n}}$, $\psi = \widehat{\mathbf{n}\,\mathbf{i}_1}$ are illustrated in Fig. 3.3.

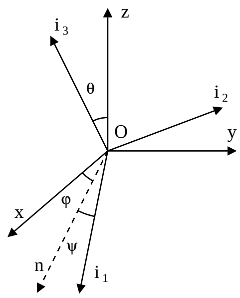


Fig. 3.3

Fig.3.3.The Euler angles.

The components in $(0; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ of $\mathbf{k}, \mathbf{i}_3, \mathbf{n}$ are, respectively:

$$\begin{aligned} \mathbf{k} &= (\sin\theta \sin\psi,\ \sin\theta\cos\psi, \cos\theta), \\ \mathbf{i}_3 &= (0, 0, 1), \\ \mathbf{n} &= (\cos\psi, \sin\psi, 0) \end{aligned} \tag{3.9.3}$$

and will be useful in the following.

To obtain a local system of regular coordinates near $\boldsymbol{\xi}_0$, remove from the $3N + 6$ redundant coordinates, just introduced, six among them by imposing the following six restrictions:

$$\sum_{i=1}^{N} m_i \boldsymbol{\kappa}^{(i)} = \mathbf{0}, \tag{3.9.4}$$

$$\sum_{i=1}^{N} (P_i - G) \wedge m_i\, \boldsymbol{\kappa}^{(i)} = \mathbf{0}, \tag{3.9.5}$$

which signify that G is actually the baricenter of $\boldsymbol{\xi}$ as well as that of $\boldsymbol{\xi}'$ and that the configuration $\boldsymbol{\xi}' \in \Sigma$ is so chosen that the system $(m_i\boldsymbol{\kappa}^{(i)})_{i=1}^N$ of $\boldsymbol{\xi}' \in \Sigma$ ("quantities of deviation") have a vanishing "angular momentum". The above restrictions should be thought of as restrictions on the choice of the reference configuration $\boldsymbol{\xi}' \in \Sigma$ a priori arbitrary.

The six coordinates that can be eliminated via Eq. (3.9.4), and Eq. (3.9.5) are, for instance, the first two components of $\boldsymbol{\kappa}^{(1)}$ in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, the three components of $\boldsymbol{\kappa}^{(2)}$, and the first component of $\boldsymbol{\kappa}^{(3)}$ still in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. The "free coordinates" $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{3N}$ will then be (orderly enumerated):

$$(\kappa_3^{(1)}, \kappa_2^{(3)}, \kappa_3^{(3)}, \kappa_1^{(4)}, \kappa_2^{(4)}, \kappa_3^{(4)}, \ldots, \theta, \varphi, \psi, (\xi_G)_1, (\xi_G)_2, (\xi_G)_3,$$

where θ, φ, ψ are the Euler angles of $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to $(\mathbf{i}, \mathbf{j}, \mathbf{k})$, while $\kappa_j^{(i)}$ are the components in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ of the deviations $\boldsymbol{\kappa}^{(i)}$.

Given the $3N$ coordinates $\boldsymbol{\beta}$, the configuration $\boldsymbol{\xi} = \boldsymbol{\Xi}(\boldsymbol{\beta})$ is built as follows:

(i) $\boldsymbol{\xi}_G = (\beta_{3N-2}, \beta_{3N-1}, \beta_{3N}) \equiv \boldsymbol{\beta}_G$ determines the baricenter G.
(ii) $\boldsymbol{\beta}_{\text{rot}} = (\beta_{3N-5}, \beta_{3N-4}, \beta_{3N-3}) \equiv (\theta, \varphi, \psi)$ determine the orientation of the axes $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$. Therefore, the positions $P_1, \ldots, P_N$ of the N points of the auxiliary configuration, called $\boldsymbol{\xi}'$ above, are determined.
(iii) The coordinates $\boldsymbol{\beta}_V = (\beta_1, \ldots, \beta_{3N-6})$ determine $(\boldsymbol{\kappa}^{(4)}, \ldots, \boldsymbol{\kappa}^{(N)})$ and, hence, the positions in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ of the points labeled $4, 5, ..., N$ and, furthermore, the coordinates $\kappa_3^{(1)}$ and $\kappa_2^{(3)}, \kappa_3^{(3)}$ of $\boldsymbol{\kappa}^{(1)}, \boldsymbol{\kappa}^{(3)}$.
(iv) The coordinates of $\boldsymbol{\kappa}^{(2)}$, as well as the remaining coordinates of $\boldsymbol{\kappa}^{(1)}, \boldsymbol{\kappa}^{(3)}$, are determined from Eq. (3.9.4) and (3.9.5). Eq. (3.9.4) yields

$$\boldsymbol{\kappa}^{(2)} = -\frac{m_1}{m_2}\boldsymbol{\kappa}^{(1)} - \sum_{i=3}^{N} \frac{m_i}{m_2}\boldsymbol{\kappa}^{(i)} \tag{3.9.6}$$

which, inserted into Eq. (3.9.5), yields

$$m_1(P_1 - P_2) \wedge \boldsymbol{\kappa}^{(1)} + \sum_{i=3}^{N} m_i(P_i - P_2) \wedge \boldsymbol{\kappa}^{(i)} = \mathbf{0} \tag{3.9.7}$$

By scalar multiplication of Eq. (3.9.7) by $(P_1 - P_2)$, it is

$$\sum_{i=3}^{N} m_i \, (P_1 - P_2) \wedge \boldsymbol{\kappa}^{(i)} \cdot (P_i - P_2) = 0 \tag{3.9.8}$$

which determines the value of $\boldsymbol{\kappa}_1^{(3)}$. In fact, recalling that $\mathbf{i}_3$ is orthogonal to the plane $\mathbf{i}_1, \mathbf{i}_2$ and that the latter three points are not aligned, $(P_3 - P_2) \wedge \mathbf{i}_1 \cdot (P_1 - P_2) \neq 0$ so that Eq. (3.9.8) is a linear equation for $\boldsymbol{\kappa}_1^{(3)}$ (with non-zero coefficient in front of $\boldsymbol{\kappa}_1^{(3)}$).

Once $\boldsymbol{\kappa}^{(3)}$ is completely determined, Eq. (3.9.7) unambiguously provides the first two components of $\boldsymbol{\kappa}^{(1)}$, because Eq. (3.9.7) only leaves the component

of $\boldsymbol{\kappa}^{(1)}$) along $(P_1 - P_2)$ undetermined, which, however, is just $\boldsymbol{\kappa}_3^{(1)}$) (recall that $\mathbf{i}_3$ is parallel, by construction, to $(P_1 - P_2)$, i.e., it is already known to be β_1).

Finally, once $\boldsymbol{\kappa}^{(1)}$ and $\boldsymbol{\kappa}^{(3)}$ are completely known, $\boldsymbol{\kappa}^{(2)}$ is derived from Eq. (3.9.5).

It is now possible to check the invertibility, near $\boldsymbol{\xi}_0$, of the transformation associating with $\boldsymbol{\beta} = (\boldsymbol{\beta}_V, \boldsymbol{\beta}_{\rm rot}, \boldsymbol{\beta}_G)$ the configuration $\boldsymbol{\Xi}(\boldsymbol{\beta}) = \boldsymbol{\xi}$ built following rules (i)-(iv). Such a transformation is also of class C^∞ with non vanishing Jacobian matrix near $\boldsymbol{\xi}_0$. However, we do not enter into the laborious analysis of the check of the regularity, invertibility, and non singularity of $\boldsymbol{\Xi}$: it does not present any conceptual problem.

Hence, the transformation $\boldsymbol{\Xi}$ establishes a regular system of local coordinates in some small enough neighborhood U of $\boldsymbol{\xi}_0 \in \Sigma$.

Clearly, the points in $\Sigma \cap U$ are those described by $\beta_1 = \ldots = \beta_{3N-6} = 0$; i.e., $(U, \boldsymbol{\Xi})$ is adapted to Σ. Actually, $(U, \boldsymbol{\Xi})$ is well adapted and orthogonal on Σ, with respect to the scalar product of Eq. (3.7.1). To show this, try to find the kinetic matrix associated with $(U, \boldsymbol{\Xi})$ and Eq. (3.7.1). For this purpose the kinetic energy of a motion $t \to \mathbf{x}(t)$ of N points has to be expressed through in terms motion $t \to \mathbf{b}(t) = (\mathbf{b}_V(t), \mathbf{b}_{\rm rot}, \mathbf{b}_G(t)) \stackrel{def}{=} \boldsymbol{\Xi}^{-1}(\mathbf{x}(t))$, assuming that the motion $\mathbf{x}$ takes place inside U for $t \in [t_1, t_2]$.

By the definition of the coordinates $\boldsymbol{\beta}$, one has, for $t \in [t_1, t_2]$:

$$\mathbf{x}^{(i)}(t) = \mathbf{b}_G(t) + \sum_{\ell=1}^{3} \big(\boldsymbol{\kappa}_\ell^{(i)}(t) + (P_i - G)_\ell \, \mathbf{i}_\ell(t)\big) \tag{3.9.9}$$

and by differentiation one finds

$$\dot{\mathbf{x}}^{(i)}(t) = \dot{\mathbf{b}}_G(t) + \sum_{\ell=1}^{3} \Big(\dot{\boldsymbol{\kappa}}^{(i)}(t) \; + \big(\boldsymbol{\kappa}^{(i)}(t) + (P_i - G)_\ell\big) \frac{d\mathbf{i}_\ell(t)}{dt}\Big) \tag{3.9.10}$$

We will now use a kinematic formula giving a simple expression to the time derivative of three mutually orthogonal unit vectors which are time dependent:

$$\frac{d\; i_\ell(t)}{dt} = \boldsymbol{\omega} \wedge \mathbf{i}_\ell(t) \tag{3.9.11}$$

here $\boldsymbol{\omega} = \boldsymbol{\omega}(t)$ is a suitable vector called "angular velocity" of the triple $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$.[10]

---

[10] To understand Eq. (3.9.11), note that, in general, the space orientation of three mutually orthogonal axes $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ imagined as emerging from a fixed point Ω can only vary if its three Euler angles (θ, φ, ψ) with respect to a fixed triple $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ change. If only θ varies, it means that the reference frame $(\Omega; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ rotates around the node line (see Fig.3.3), and it is then clear a every point P co-moving with $(\Omega; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ has velocity $\mathbf{v}_P = \dot\theta \, \mathbf{n} \cdot (P - \Omega)$. This holds, in particular, for the extremities of $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$; hence:

A useful expression for $\boldsymbol{\omega}$ is, in terms of the Euler angles (see p.202, footnote 10):

$$\boldsymbol{\omega} = \dot{\theta} + \dot{\varphi}\mathbf{k} + \dot{\psi}\mathbf{i}_3 \tag{3.9.12}$$

Coming back to Eq. (3.9.10), we shall rewrite it by using Eq. (3.9.1)

$$\dot{\mathbf{x}}(t) = \dot{\mathbf{b}}_G + \dot{\tilde{\boldsymbol{\beta}}}^{(i)} + \boldsymbol{\omega} \wedge (\boldsymbol{\kappa}^{(i)} + (P_i - G)), \tag{3.9.13}$$

$$\frac{d\mathbf{i}_\ell}{dt} = \dot{\theta}\,\mathbf{n} \wedge \mathbf{i}_\ell, \qquad \ell = 1, 2, 3,$$

A similar argument shows that if $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ move because only φ or ψ vary, then

$$\frac{d\mathbf{i}_\ell}{dt} = \dot{\varphi}\,\mathbf{k} \wedge \mathbf{i}_\ell, \quad \text{or} \quad \frac{d\mathbf{i}_\ell}{dt} = \dot{\psi}\,\mathbf{i}_3 \wedge \mathbf{i}_\ell,$$

More generally, if $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ vary because θ, φ, ψ simultaneously vary, it will be (by the differentiation rule of composed functions):

$$\frac{d\mathbf{i}_\ell}{dt} = \boldsymbol{\omega} \wedge \mathbf{i}_\ell, \qquad \ell = 1, 2, 3$$

with $\boldsymbol{\omega}$ given by $\dot{\theta}\mathbf{n} + \dot{\varphi}\mathbf{k} + \dot{\psi}\mathbf{i}_3$, i.e., Eq. (3.9.11).

In connection with Eq. (3.9.11), it is natural to note one of its consequences: the relation between a motion $t \to P(t)$ in a frame $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$ and the same motion in a frame $(\Omega(t); \mathbf{i}_1(t), \mathbf{i}_2(t), \mathbf{i}_3(t))$, time dependent. From the vector relation

$$P(t) - O = (P(t) - \Omega(t)) + (\Omega(t) - O)$$

written componentwise as

$$x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} = x_1(t)\mathbf{i}_1(t) + x_2(t)\mathbf{i}_2(t) + x_3(t)\mathbf{i}_3(t) + x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k},$$

with obvious notations, it follows, by differentiation, that

$$\mathbf{V}^{(a)} = \mathbf{V}^{(r)} + x_1\frac{d\mathbf{i}_1}{dt} + x_2\frac{d\mathbf{i}_2}{dt} + x_3\frac{d\mathbf{i}_3}{dt} + \mathbf{V}_\Omega, .$$

where $V^{(a)} = \dot{x}(t)\mathbf{i} + \dot{y}(t)\mathbf{j} + \dot{z}(t)\mathbf{k}$ is the velocity of $t \to P(t)$ in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, $\mathbf{V}^{(r)}$ is the velocity of same motion "relative" to $(\Omega(t); \mathbf{i}_1(t), \mathbf{i}_2(t), \mathbf{i}_3(t))$, i.e., $\mathbf{V}^{(r)} = \dot{x}_1(t)\mathbf{i}_1(t) + \dot{x}_2(t)\mathbf{i}_2(t) + \dot{x}_3(t)\mathbf{i}_3(t)$ and $\mathbf{V}_\Omega$ is the velocity of the motion $t \to \Omega(t)$ in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, i.e., $\mathbf{V}_\Omega = \dot{x}(t)\mathbf{i} + \dot{y}(t)\mathbf{j} + \dot{z}(t)\mathbf{k}$. Then, by using Eq. (3.9.11):

$$\mathbf{V}^{(a)} = \mathbf{V}^{(r)} + \boldsymbol{\omega}\wedge(x_1(t)\mathbf{i}_1(t) + x_2(t)\mathbf{i}_2(t) + x_3(t)\mathbf{i}_3(t)) + \mathbf{V}_\Omega = \mathbf{V}^{(r)} + (\boldsymbol{\omega}\wedge(P - \Omega) + \mathbf{V}_\Omega).$$

The term in parentheses has the interpretation of the "drag velocity" that the point P would have if it were fixed in $(\Omega; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$; hence, the above formula reads "the absolute speed equals the sum of the relative speed plus the drag speed". Furthermore, the velocity of a point P *fixed* in a moving frame $(\Omega; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ is given by

$$\mathbf{V}_P = \mathbf{V}_\Omega + \boldsymbol{\omega} \wedge (P - \Omega),$$

where $\mathbf{V}_P$ is the velocity in $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$ of P, $\boldsymbol{\omega}$ is the angular velocity of the triplet $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, and V_Ω is the speed of Ω in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$. The last relation is of great interest in the theory of rigid motion.

where we set $\dot{\tilde{\boldsymbol{\kappa}}}^{(i)} = \sum_{\ell=1}^{3} \dot{\boldsymbol{\kappa}}_\ell^{(i)} \mathbf{i}_\ell(t)$ (which differs from $\dot{\boldsymbol{\kappa}}^{(i)}(t)$; in fact, it is the velocity of the i-th point relative to the moving frame, while $\dot{\boldsymbol{\kappa}}^{(i)}$ is its velocity relative to the fixed frame), and $(P_i - G) = \sum_{\ell=1}^{3}(P_i - G)_\ell \mathbf{i}_\ell(t)$.

It is now possible to compute the kinetic energy, using Eq. (3.9.13):

$$
\begin{aligned}
T = \frac{1}{2}\sum_{i=1}^{N} m_i(\dot{\mathbf{x}}^{(i)})^2 &= \frac{1}{2}\sum_{i=1}^{N} m_i\big(\dot{\mathbf{b}}_G^2 + \boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)} + (P_i - G)) + \dot{\tilde{\boldsymbol{\kappa}}}^{(i)}\big)^2 \\
&= \frac{1}{2}(\sum_{i=1}^{N} m_i)\dot{\mathbf{b}}_G^2 + \frac{1}{2}\sum_{i=1}^{N} m_i(\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)} + (P_i - G)))^2 \\
&\quad + \frac{1}{2}\sum_{i=1}^{N} m_i(\dot{\tilde{\boldsymbol{\kappa}}}^{(i)})^2 + \sum_{i=1}^{N} m_i\dot{\mathbf{b}}_G\cdot\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)} + (P_i - G)) \\
&\quad + \sum_{i=1}^{N} m_i\dot{\mathbf{b}}_G\cdot\dot{\tilde{\boldsymbol{\kappa}}}^{(i)} + \sum_{i=1}^{N} m_i\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)} + (P_i - G))\cdot\dot{\tilde{\boldsymbol{\kappa}}}^{(i)}
\end{aligned}
\tag{3.9.14}
$$

The fourth and fifth terms in the right-hand side vanish identically: which follows by taking the constant vectors out of the summations and recalling the definition of the baricenter (by which $\sum_{i=1}^{N} m_i(P_i - G) = \mathbf{0}$ as well as Eq. (3.9.4). To study the second and the sixth terms of the right-hand side (3.9.14), we will use the formula

$$
(\mathbf{a}\wedge\mathbf{b})\cdot\mathbf{c} = (\mathbf{b}\wedge\mathbf{c})\cdot\mathbf{a} = (\mathbf{c}\wedge\mathbf{a})\cdot\mathbf{b} \tag{3.9.15}
$$

to note that

$$
\sum_{i=1}^{N} m_i\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)} + (P_i - G))\cdot\dot{\tilde{\boldsymbol{\kappa}}}^{(i)} = \boldsymbol{\omega}\cdot\big(\sum_{i=1}^{N} m_i(\boldsymbol{\kappa}^{(i)} + (P_i - G))\wedge\dot{\tilde{\boldsymbol{\kappa}}}^{(i)}\big), \tag{3.9.16}
$$

and one can remark that the quantity within brackets in the right-hand of Eq. (3.9.16) is the angular momentum $\mathbf{K}_G^{(\mathrm{in})}$ "relative" to the frame $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ (also called the "internal angular momentum") and, furthermore, by Eq. (3.9.5) written componentwise in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, by the $\dot{\tilde{\boldsymbol{\kappa}}}^{(i)}$ definition and by the time independence of the components of $(P_i - G)$ in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, it follows that $\sum_{i=1}^{N} m_i(P_i - G)) \wedge \dot{\tilde{\boldsymbol{\kappa}}}^{(i)} = \mathbf{0}$,[11] so that

$$
\mathbf{K}_G^{(\mathrm{in})} = \sum_{i=1}^{N} m_i(\boldsymbol{\kappa}^{(i)} + (P_i - G))\wedge\dot{\tilde{\boldsymbol{\kappa}}}^{(i)} \equiv \sum_{i=1}^{N} m_i\,\boldsymbol{\kappa}^{(i)}\wedge\dot{\tilde{\boldsymbol{\kappa}}}^{(i)} \tag{3.9.17}
$$

[11] Let $s = 1, 2, 3$ then Eq. (3.9.5) gives

it is therefore true and, as it will be seen, important that $\mathbf{K}_G^{(\mathrm{in})} = \mathbf{0}$ if $\boldsymbol{\kappa}^{(1)} = \ldots = \boldsymbol{\kappa}^{(N)} = \mathbf{0}$, i.e., if the system is, at the time considered, on Σ. The second term of the right-hand side of Eq. (3.9.14) will be written as

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{N} m_i\big(\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)}+(P_i-G))\big)^2 \\ &= \frac{1}{2}\boldsymbol{\omega}\cdot\big[\sum_{i=1}^{N} m_i(\boldsymbol{\kappa}^{(i)}+(P_i-G))\wedge(\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)}+(P_i-G)))\big] \\ &= \frac{1}{2}(\boldsymbol{\omega}\cdot\mathbf{I}(\boldsymbol{\omega})), \end{aligned} \tag{3.9.18}$$

where Eq. (3.9.15) has been used, having defined

$$\mathbf{I}(\boldsymbol{\omega}) \stackrel{def}{=} \sum_{i=1}^{N} m_i(\boldsymbol{\kappa}^{(i)}+(P_i-G))\wedge(\boldsymbol{\omega}\wedge(\boldsymbol{\kappa}^{(i)}+(P_i-G))). \tag{3.9.19}$$

Then define

$$T_G \stackrel{def}{=} \frac{1}{2}(\sum_{i=1}^{N} m_i)\dot{\mathbf{x}}_G^2, \qquad \text{“baricenter kinetic energy”}, \tag{3.9.20}$$

$$T^{(\mathrm{in})} \stackrel{def}{=} \frac{1}{2}\sum_{i=1}^{N} m_i(\dot{\tilde{\boldsymbol{\kappa}}}^{(i)})^2, \qquad \text{“internal kinetic energy”}, \tag{3.9.21}$$

$$T_C \stackrel{def}{=} \mathbf{K}_G^{(\mathrm{in})}\cdot\boldsymbol{\omega}, \quad \text{“complementary” or “Coriolis” kinetic energy}, \tag{3.9.22}$$

$$T_{\mathrm{rot}} \stackrel{def}{=} \frac{1}{2}\boldsymbol{\omega}\cdot\mathbf{I}(\boldsymbol{\omega}), \qquad \text{“rotational kinetic energy”}, \tag{3.9.23}$$

and remark that it has just been shown that

$$T = T_G + T^{(\mathrm{in})} + T_{\mathrm{rot}} + T_C. \tag{3.9.24}$$

When $\boldsymbol{\omega} = \mathbf{0}$, this relation is called “König's theorem”.

From Eq. (3.9.24), it can be seen that the coordinate system defined by $\boldsymbol{\Xi}$ near $\boldsymbol{\xi}_0$ is well adapted and orthogonal on Σ. In fact, one can note that at a

---

$$\begin{aligned} 0 =& \frac{d}{dt}\sum_{i=1}^{N} m_i\Big((\boldsymbol{\kappa}^{(i)}+(P_i-G))\wedge\boldsymbol{\kappa}^{(i)}\Big)_s = \frac{d}{dt}\sum_{i=1}^{N} m_i \sum_{\ell',\ell''}^{1,3}(P_i-G)_{\ell'}\kappa_{\ell''}^{(i)}(\mathbf{i}_{\ell'}\wedge\mathbf{i}_{\ell''})_s \\ =& \sum_{i=1}^{N} m_i \sum_{\ell',\ell''}^{1,3}(P_i-G)_{\ell'}\dot{\kappa}_{\ell''}^{(i)}(\mathbf{i}_{\ell'}\wedge\mathbf{i}_{\ell''})_s = \Big(\sum_{i=1}^{N} m_i(P_i-G)\wedge\dot{\tilde{\boldsymbol{\kappa}}}^{(i)}\Big)_s \end{aligned}$$

since $(\mathbf{i}_{\ell'}\wedge\mathbf{i}_{\ell''})_s$ is either 0 or ±1, $\forall t$, i.e. it has 0 t-derivative.

point $\boldsymbol{\xi} \in \Sigma$, the coordinates $\boldsymbol{\kappa}^{(1)}, \dots, \boldsymbol{\kappa}^{(N)}$, hence, the $\boldsymbol{\beta}_V$'s, vanish (i.e., the $\boldsymbol{\beta}$ coordinates are adapted to Σ). Furthermore, if the motion $t \to \mathbf{x}(t)$ happens to occupy the position $\boldsymbol{\xi} \in \Sigma$ at a certain time t_0 we have $T_C(t_0) = 0$, by Eqs. (3.9.17) and (3.9.22).

At the same time instant $T^{(\mathrm{in})}(t_0)$, as one realizes from the determination of $\boldsymbol{\kappa}^{(1)}, \boldsymbol{\kappa}^{(2)}, \boldsymbol{\kappa}^{(3)}$, via Eqs. (3.9.4) and (3.9.5), is a quadratic form in $\dot{\mathbf{b}}_V(t_0)$ with coefficients only depending upon the structure of Σ via the coordinates $(P_i - G)_\ell$, $i = 1, 2, 3$, $\ell = 1, \dots, N$, which are given constants ($\boldsymbol{\xi}$ independent, hence, $\boldsymbol{\beta}_V$ independent).

Finally, $T_G(t_0)$ is a quadratic form in $\dot{\mathbf{b}}_G(t_0)$ with constant coefficients, while $T_{\mathrm{rot}}(t_0)$ is a quadratic form in the components of $\boldsymbol{\omega}$ in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with coefficients depending only on the structure of Σ via the (constant) coordinates $(P_i - G)_\ell, \ell = 1, 2, 3$, $i = 1, \dots, N$ [see, for more details, Eq. (3.9.29)]. Hence, since the components of $\boldsymbol{\omega}$ in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ are, by Eq. (3.9.3),

$$\begin{aligned}\omega_1 &= \dot{\theta} \cos \psi + \dot{\varphi} \sin \theta \sin \psi \\ \omega_2 &= -\dot{\theta} \sin \psi + \dot{\varphi} \sin \theta \cos \psi \\ \omega_3 &= \dot{\varphi} \cos \theta + \dot{\psi}\end{aligned} \tag{3.9.25}$$

it follows that the rotation kinetic energy is a quadratic form in $\dot{\mathbf{b}}_{\mathrm{rot}}(t_0) = (\dot{\theta}, \dot{\varphi}, \dot{\psi})_{t=t_0}$ with coefficients solely dependent on θ, φ, ψ [by Eq. (3.9.25)].

Hence, the quadratic forms defining T on Σ do not contain any mixed terms like or $(\dot{\mathbf{b}}_V)_i (\dot{\mathbf{b}}_{\mathrm{rot}})_j$ or $(\dot{\mathbf{b}}_V)_i (\dot{\mathbf{b}}_G)_j$; therefore, the coordinate system is orthogonal on Σ (see Definition 12, §3.7, p.177). It is also well adapted by the above observed constancy of the coefficients of the quadratic form in $\dot{\mathbf{b}}_V(t_0)$ expressing $T^{(\mathrm{in})}(t_0)$.

From the definition of W, it appears that W depends only upon $\boldsymbol{\kappa}^{(1)}, \dots,$ $\boldsymbol{\kappa}^{(N)}$ through their components in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, i.e., only upon $\boldsymbol{\beta}_V$ [in fact, as already remarked, such components can be reconstructed from the $\boldsymbol{\beta}_V$'s via Eqs. (3.9.4) and (3.9.5) and depend only on the $\boldsymbol{\beta}_V$'s and do not depend on $(\boldsymbol{\beta}_{\mathrm{rot}}, \boldsymbol{\beta}_G)$].

This concludes the perfection proof for the constraint model (Σ, W, λ) in the rigid case considered above. mbe

Observation. By deducing Eq. (3.9.24), it has been explicitly shown that the kinetic energy of a rigid body, i.e., of a motion of a system of N masses in $\mathcal{R}^3$ constrained to keep fixed mutual distances, can be expressed in terms of six coordinates and their derivatives. If such coordinates are the baricenter coordinates $\mathbf{x}_G$ and the three Euler angles (θ, φ, ψ) of a co-moving frame $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to a fixed frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ and if $\omega_1, \omega_2, \omega_3$ are the components in $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ of the angular velocity [see Eqs. (3.9.12), (3.9.3), and (3.9.25), then there exists a 3×3 matrix $I = (I_{ij})_{i,j=1,2,3}$ such that

$$T = \frac{1}{2} M \dot{\mathbf{x}}_G^2 + \frac{1}{2} \sum_{i,j=1}^{3} I_{ij} \omega_i \omega_j, \tag{3.9.26}$$

where $M = \sum_{i=1}^{N} m_i$ and T denotes the system kinetic energy. In fact, Eq. (3.9.26) follows from Eq. (3.9.24), since in this case $\boldsymbol{\kappa}^{(i)} = \mathbf{0}$ (because the motion is rigid) and $T^{(\text{in})} \equiv 0 \equiv T_C$, and from Eq. (3.9.18) showing

$$\begin{aligned} T_{\text{rot}} =& \frac{1}{2}\boldsymbol{\omega} \cdot \mathbf{I}(\boldsymbol{\omega}) = \frac{1}{2}\sum_{i=1}^{N} m_i \big[(P_i - G) \wedge (\boldsymbol{\omega} \wedge (P_i - G))\big] \cdot \boldsymbol{\omega} \\ =& \frac{1}{2}\sum_{i=1}^{N} m_i (\boldsymbol{\omega} \wedge (P_i - G))^2 \end{aligned} \tag{3.9.27}$$

by Eq. (3.9.15); then Eq. (3.9.27) permits us to obtain Eq. (3.9.26) as follows. If θ_i is the angle between $\boldsymbol{\omega}$ and $(P_i - G)$:

$$\begin{aligned} &(\boldsymbol{\omega} \wedge (P_i - G))^2 = \boldsymbol{\omega}^2 \,(P_i - G)^2 \,(\sin\theta_i)^2 = \boldsymbol{\omega}^2 \,(P_i - G)^2 \,(1 - (\cos\theta_i)^2) \\ &= \boldsymbol{\omega}^2 \,(P_i - G)^2 \,\big[1 - \frac{(\boldsymbol{\omega} \cdot (P_i - G))^2}{\boldsymbol{\omega}^2 \,(P_i - G)^2}\big] = \boldsymbol{\omega}^2 \,(P_i - G)^2 - (\boldsymbol{\omega} \cdot (P_i - G))^2 \end{aligned}$$

$$\begin{aligned} &= (\sum_{\ell=1}^{3} \omega_\ell^2)\,(\sum_{\ell'=1}^{3} (P_i - G)_{\ell'}^2) - \sum_{\ell,\ell'=1}^{3} \omega_\ell \omega_{\ell'} (P_i - G)_\ell (P_i - G)_{\ell'} \\ &\equiv \sum_{\ell,\ell'=1}^{3} \omega_\ell \omega_{\ell'} \big((\sum_{\widetilde{\ell}=1}^{3} (P_i - G)_{\widetilde{\ell}}^2)\delta_{\ell\ell'} - (P_i - G)_\ell (P_i - G)_{\ell'}\big) \end{aligned} \tag{3.9.28}$$

hence,

$$I_{\ell\ell'} = \sum_{i=1}^{N} m_i \big\{(\sum_{\widetilde{\ell}=1}^{3} (P_i - G)_{\widetilde{\ell}}^2)\delta_{\ell\ell'} - (P_i - G)_\ell (P_i - G)_{\ell'}\big\}, \tag{3.9.29}$$

which are constants, $\forall \ell, \ell' = 1, 2, 3$, characteristic of the rigid body because such are the components $(P_i - G)_\ell, \ell = 1, \ldots, N, \ell = 1, 2, 3$, of the vectors $(P_i - G)$ in the co-moving frame $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. We shall come back to Eqs. (3.9.26) and (3.9.29), deducing them independently of the constraint theory, to help the readers who have not paid attention to the proofs of this section.

3.9.1 Exercises and Problems

1. Suppose that the reference system $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, with origin at the baricenter of a system of masses, has a purely translational motion in the reference system $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$. Show that the kinetic energy is $T = T_G + T^{(\text{in})}$.

2. Let $t \to \omega(t), t \in \mathcal{R}_+$, be the angular velocity for the triplet of orthogonal unit vectors $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ moving in the reference frame $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$. Let $t \to \mathbf{y}(t)$, $t \in \mathcal{R}_+$, be a $\mathcal{R}^3$-valued

function and write it as $\mathbf{y}(t) = \sum_{\ell=1}^{3} y_\ell(t)\mathbf{i}_\ell(t)$. Define $\dot{\tilde{\mathbf{y}}}(t) = \sum_{\ell=1}^{3} \dot{y}_\ell(t)\mathbf{i}_\ell(t)$ and show that $\dot{\mathbf{y}} = \dot{\tilde{\mathbf{y}}} + \boldsymbol{\omega} \wedge \mathbf{y}$ by using $\frac{d\mathbf{i}_\ell}{dt} = \boldsymbol{\omega} \wedge \mathbf{i}_\ell$. Show also that $\dot{\boldsymbol{\omega}} = \dot{\tilde{\boldsymbol{\omega}}}$.

3. Compute the components of $\boldsymbol{\omega}$, Eq. (3.9.12), in $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$ in terms of the Euler angles and their derivatives.

4. Evaluate the matrix $I_{j\ell}$ for the rigid system in Fig. 3.2(a), assuming $m_0 = m_1 = m_2 = m_3 = 1$ or $m_0 = 1$, $m_1 = 2$, $m_2 = 3$, $m_3 = 4$, taking for the moving frame the one with axes parallel to those in Fig. 3.2(a) and origin in G (*Hint:* If the direct computation looks cumbersome, replace G by O using Problems 5 and 6 below.)

5. Consider a rigid system constrained to have one of its points fixed at the origin of the fixed frame of reference $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$. Show that if $(O; \mathbf{i}_1, \mathbf{i}_3, \mathbf{i}_3)$ is a co-moving frame, the kinetic energy can be expressed as $T = \frac{1}{2} \sum_{\ell,\ell'} J_{\ell\ell'} \omega_\ell \omega_{\ell'}$, where $J_{\ell\ell'}$ are constants depending only on the body structure.

6. In the context of Problem 5, consider the cases when $O = G$ and when $O \neq G$, calling $I_{\ell\ell'}$ or $J_{\ell\ell'}$ the matrix expressing the kinetic energy in the frames $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ or $(G; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. Show that if $M = \sum_{i=1}^{N} m_i$,

$$J_{\ell\ell'} = I_{\ell\ell'} + M[(G-O)^2 \delta_{\ell\ell'} - (G-O)_\ell (G-O)_{\ell'}].$$

3.10 General Considerations on the Theory of Constraints

The approximate constraint theory, in the analysis of Proposition 13, §3.8, and Proposition 14, §3.9, still contains some unsatisfactory aspects that it is useful to mention explicitly.

In applications in which a certain model (Σ, W, λ) of approximately ideal constraint is a good model, the rigidity parameter λ has a well-defined value $\lambda < +\infty$ which is fixed and, therefore, cannot tend to $+\infty$.

Therefore, the problem arises of how to estimate, in terms of λ, the error encountered when approximating the "real motions" $t \to \mathbf{x}_\lambda(t)$ with their limits as $\lambda \to +\infty$ (which are described by the equations of motion relative to ideal constraints, because (Σ, W, λ) is supposed to be an approximately ideal constraint, i.e., by "simple" equations).

The theory of §3.8, if one carefully looks at the formulas derived in the proof, also provides some estimates of the errors made in the mentioned approximation.

However, it is sufficient to simply look at the calculations made there to realize that if N is a number of the order of magnitude of a few dozens (not to speak of the cases when it is on the order of Avogadro's number, as is sometimes the case), such estimates become ridiculously rough for reasonable values of λ and reasonable models of W.

As usual, the problem of finding good error estimates is a problem that should not be posed in too great a generality but should be discussed in connection with precise and concrete questions of a physical nature concerning

the behavior of physical entities which, in each case, appear as interesting. Even so, it remains a very difficult question and is a typical problem in statistical mechanics. Except in a few simple cases, it is an essentially open problem from a mathematical viewpoint.

Physicists and engineers have elaborated theories, mathematically non rigorous consequences of the dynamics of point masses, which allow them to evaluate the errors involved in the perfect constraint approximation in a reasonable way, often experimentally correct.[12]

However, it is often only through recourse to experiments that one is able to understand whether a certain constraint can or cannot be approximated by an ideal one.

It is good for the student to keep the above considerations in mind while solving the standard book-made problems concerning the constrained motions in order to appreciate their often purely didactic and abstract nature.

The above discussion, which we will not continue, gives an idea of the depth of the ideal constraint notion, and it can perhaps be useful to understand why long and learned discussions on the argument often take place. So many and so diverse are these arguments that they may leave those who realize their existence for the first time quite surprised.

Other problems naturally arise in the theory of the holonomous constraints. Some of them are:

(i) When an approximate constraint (Σ, W, λ) is not perfect, how can the motion be described in the limit $\lambda \to +\infty$? Is it possible, as the considerations in §3.6 seem to suggest, to treat the constraint in this limit as ideal in the sense of §3.5, modifying the potential energy of the active forces, possibly as a function of the initial datum? See the example of §3.6, following Definition 10.

(ii) In case (i), how can we find the active forces? And how can we estimate the errors involved in the approximation $\lambda = +\infty$?

(iii) How can we treat the case when the constraint model (Σ, W, λ) is ideal but the system moves under the influence of a force which is the sum of the constraint force, with potential energy λW, and a force law, in $C^\infty(\mathcal{R}^{Nd})$

$$\mathbf{F}^{(a)} = \mathbf{F}^{(a)}(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(N)}) \tag{3.10.1}$$

which is not necessarily conservative?

(iv) This is the same as (iii), replacing the constraint model by an approximate conservative model which is not approximately ideal.

(v) This is the same as (i) and (ii) in the situation described by (iii).

[12] For instance, elasticity theory has, among other theories, this scope. Of course, elasticity theory can be set up as a mathematically rigorous theory in itself: what is non rigorous is the connection between elasticity theory and the above microscopic theory of constraints. In other words, elasticity theory is itself a mathematical model which in this case "models" another mathematical model: even such things can happen!

The preceding problems are not easy and are open problems to some extent (in the sense that there do not seem to be in the literature any interesting general propositions about them) except problem (iii) which is essentially completely solved by the following proposition, proved exactly in the same way as the analogous Proposition 13, §3.8, p.186:

15 Proposition. *Let (Σ, W, λ) be a model for an ideal approximate bilateral s-codimensional constraint for N points in $\mathcal{R}^d$, with masses $m_1, \ldots, m_N > 0$. Consider an initial datum $(\boldsymbol{\eta}_0, \boldsymbol{\xi}_0) \in \mathcal{R}^{2Nd}$ such that $\boldsymbol{\xi}_0 \in \Sigma$. Let $t \to \mathbf{x}(t)$ be the motion that follows this initial datum and develops under the influence of the field of conservative forces with potential energy λW and of a field $\mathbf{F}^{(a)} \in C^\infty_{\text{lim}}(\mathcal{R}^{Nd})$ of uniformly bounded forces, not necessarily conservative. Then the limit*

$$\lim_{\lambda \to +\infty} \mathbf{x}_\lambda(t) = \mathbf{x}(t) \tag{3.10.2}$$

exists for every $t \in \mathcal{R}$ and it is a motion constrained to Σ with initial datum

$$\mathbf{x}(0) = \boldsymbol{\xi}_0, \quad \dot{\mathbf{x}}(0) = \boldsymbol{\eta}_0^\Sigma \tag{3.10.3}$$

[see Eqs. (3.7.24) and (3.7.26)]. Suppose that for $t \in [t_1, t_2]$ the motion $\mathbf{x}$ dwells in a neighborhood U where a system $(U, \boldsymbol{\Xi})$ of local regular coordinates adapted to Σ is established. Then $\mathbf{x}$ is described in the basis Ω for $(U, \boldsymbol{\Xi})$ by a motion $t \to \mathbf{b}(t)$ verifying:

$$b_1(t) = b_2(t) = \ldots = b_s(t) = 0, \tag{3.10.4}$$

$$\frac{d}{dt}\Big(\frac{\partial T}{\partial \alpha_i}(\dot{\mathbf{b}}(t), \mathbf{b}(t))\Big) - \Big(\frac{\partial T}{\partial \beta_i}(\dot{\mathbf{b}}(t), \mathbf{b}(t))\Big) = \Phi_i(\mathbf{b}(t)), \tag{3.10.5}$$

$\forall i = s+1, \ldots, Nd$, where

$$T(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{2}\sum_{i,j=1}^{Nd} g_{ij}(\boldsymbol{\beta})\alpha_i\alpha_j, \qquad (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{R}^{Nd} \times \Omega \tag{3.10.6}$$

if g is the kinetic matrix associated with the system $(U, \boldsymbol{\Xi})$ and

$$\Phi_i(\boldsymbol{\beta}) = \sum_{k=1}^{N} \mathbf{F}^{(a)(k)}(\boldsymbol{\Xi}(\boldsymbol{\beta})) \cdot \frac{\partial \boldsymbol{\Xi}^{(k)}}{\partial \beta_i}(\boldsymbol{\beta}) \tag{3.10.7}$$

Observation. The functions in Eq. (3.10.7) on Ω are called the "force components" of the force $\mathbf{F}^{(a)} = (\mathbf{F}^{(a)(1)}, \ldots, \mathbf{F}^{(a)(N)})$ in the reference system $(U, \boldsymbol{\Xi})$. The proof of Proposition 15 is a repetition of that of Proposition 13.

A final comment on the theorems of §3.8 and §3.10 concludes this section. The condition $\boldsymbol{\xi}_0 \in \Sigma$ appears to be somewhat unnatural, and one would like

to change it to "$\boldsymbol{\xi}_0$ close enough to Σ". However, the problem is what is meant by "close enough"?

It is quite clear that the closeness notion should be λ dependent: in fact, we shall call $\boldsymbol{\xi}_0$ close to Σ only if the energy $\lambda W(\boldsymbol{\xi}_0)$ is not too large; i.e., if the initial deviation out of Σ does not involve "too large constraint forces" or "too large elastic deformation energy".

It is then clear that it will be possible to try to prove propositions analogous to Proposition 13 or Proposition 15 by replacing the hypothesis $\boldsymbol{\xi}_0 \in \Sigma$ with the hypothesis that the position of the initial datum is a function of λ, $\boldsymbol{\xi}_0(\lambda)$, such that the limit $\lim_{\lambda\to+\infty}\boldsymbol{\xi}_0(\lambda) = \boldsymbol{\xi}_0 \in \Sigma$. In this case, $\lambda W(\boldsymbol{\xi}_0(\lambda))/\lambda \xrightarrow[\lambda\to+\infty]{} 0$ (i.e., the initial "constraint deformation energy" is not too large, being of lower order with respect to λ, which is the order of the energy of a λ-independent deformation).

The proof of the analogues of Propositions 13 and 15 would be identical under these more general assumptions: this could be realized via a detailed examination of their proofs.

Historical Note: The idea that the constrained systems, ideal or not, could be thought of as limiting cases of non constrained systems subject to strong forces is naturally ancient. However, to the best of this author's knowledge, it has been written down in the form of a precise theorem to be interpreted as a proof of the least-action principle in [1] (p.80-82). Here the idea is expressed and it is shown how the least-action principle can be deduced through Proposition 13, §3.8, p.186. This is, in my opinion, the most interesting and deepest of the "proofs" of the least-action principle (and, hence, of the virtual-work principle). There exist other proofs, sometimes very ingenious, which, however, are never more than pseudo-proofs in the sense well described by E. Mach ([31], e.g., in Chapter III, §5.6).

3.11 Equations of Hamilton and Lagrange. Analytical Mechanics

Before beginning the study of concrete mechanical problems, it is convenient to deduce from what has already been seen some abstract mathematical structure naturally arising in the context of constraint theory and the least-action principle.

14 Definition. *Let $U \subset \mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$ be an open set and let $\mathcal{L} \in C^\infty(U)$ be a real-valued function. $\mathcal{L}$ will be called a "regular Lagrangian function" on U if the map $\boldsymbol{\Xi}$ transforming the point $(\boldsymbol{\alpha}, \mathbf{q}, t) \in U$ into the point*

$$(\boldsymbol{\pi}, \mathbf{q}, t) \in \mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R} \tag{3.11.1}$$

with

$$\pi_i = \frac{\partial \mathcal{L}}{\partial \alpha_i}(\boldsymbol{\alpha}, \mathbf{q}, t) \tag{3.11.2}$$

maps the neighborhood U into a neighborhood $V \subset \mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$, $V = \boldsymbol{\Xi}(U)$ invertibly and with a non vanishing Jacobian ("nonsingularly").[13]
If $\mathcal{L}$ is a regular Lagrangian on U, the equations for the motion $t \to (\dot{\mathbf{q}}(t), \mathbf{q}(t), t)$ in U,

$$\frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial \alpha_i}(\dot{\mathbf{q}}, \mathbf{q}, t)\Big) = \frac{\partial \mathcal{L}}{\partial q_i}(\dot{\mathbf{q}}, \mathbf{q}, t), \tag{3.11.3}$$

are called the "Lagrangian differential equations" for the Lagrangian $\mathcal{L}$.
Since the map (3.11.1) does not really involve $\mathbf{q}$, the above definition makes sense without change if U is an open subset of $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or of $\mathcal{R}^\ell \times (\mathcal{T}^{\ell_1} \times \mathcal{R}^{\ell_2}) \times \mathcal{R}$, with $\ell_1 + \ell_2 = \ell$, provided $C^\infty(U)$ is understood in the natural sense following Definition 13, p.101, §2.21.
In these cases, V will have to be a subset of $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or of $\mathcal{R}^\ell \times (\mathcal{T}^{\ell_1} \times \mathcal{R}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$, and the points on the tori are to be thought of as described in "angular coordinates", see Definition 12, p.100, §2.21.

Observations:
(1) The usefulness of the clumsy-looking extension appearing in the second part of the definition can be understood by noting that, for instance, a point, with mass $m > 0$, bound to a vertically placed circle with radius R by an ideal constraint and subject to gravity has, if φ is the natural angular coordinate on the circle thought of as $\mathcal{T}^1$, a Lagrangian description in the sense of Definition 14 in terms of $\mathcal{L}(\alpha, \varphi, t) = \frac{1}{2} m\alpha^2 + mgR\cos\varphi$. In this case, $U = \mathcal{R} \times \mathcal{T}^1 \times \mathcal{R}$ and Eq. (3.11.3) becomes the pendulum equation (g being gravitational acceleration).

Similarly, a free particle ideally bound to a circle will be described on $\mathcal{R} \times \mathcal{T}^\ell \times \mathcal{R}$ by $\mathcal{L}_0(\alpha, \varphi, t) = \frac{1}{2} m\alpha^2$.

Hence, when the surface Σ generated by an ideal constraint is topologically a torus, we have the possibility of using "global angular coordinates" without having to cover Σ, to describe the motions on Σ, with several local systems of regular coordinates.
(2) When $\mathcal{L}$ does not depend explicitly on time, i.e., $\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \equiv \widetilde{\mathcal{L}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$, $\forall(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \in U$, for some $\widetilde{\mathcal{L}}$, we say that $\mathcal{L}$ is "time independent" and we shall write it without the variable t.

The following proposition holds.

16 Proposition. *Let $\mathcal{L}$ be a regular Lagrangian on an open subset $U \subset \mathcal{R}^\ell \times \mathcal{R}^\ell \times R$ (or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$, $\ell_i \geq 0$), and let $t \to (\dot{\mathbf{q}}(t), \mathbf{q}(t), t) \in U$ be a motion defined for $t \in [t_1, t_2]$, verifying Eq. (3.11.3). Setting*

[13] The Jacobian determinant coincides with the determinant of the matrix $J_{ij} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\alpha}, \mathbf{q}, t)}{\partial \alpha_i \partial \alpha_j}, i, j = 1, \ldots, \ell$.

$$(\boldsymbol{\alpha}(\mathbf{p},\mathbf{q},t),\mathbf{q},t) = \boldsymbol{\Xi}^{-1}(\mathbf{p},\mathbf{q},t), \tag{3.11.4}$$

$$H(\mathbf{p},\mathbf{q},t) = \sum_{i=1}^{\ell} p_i\alpha_i(\mathbf{p},\mathbf{q},t) - \mathcal{L}(\boldsymbol{\alpha}(\mathbf{p},\mathbf{q},t),\mathbf{q}(t),t) - \mathcal{L}(\boldsymbol{\alpha}(\mathbf{p},\mathbf{q},t),\mathbf{q},t), \tag{3.11.5}$$

for $(\mathbf{p},\mathbf{q},t) \in V$ *(see Definition 14), the motion in* V*, image of the preceding motion in* U *via Eq. (3.11.1),* $t \to (\mathbf{p}(t),\mathbf{q}(t),t) = \boldsymbol{\Xi}(\dot{\mathbf{q}}(t),\mathbf{q}(t),t)$*, verifies the equations:*

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}(\mathbf{p}(t),\mathbf{q}(t),t), \qquad i = 1,\dots,\ell \tag{3.11.6}$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i}(\mathbf{p}(t),\mathbf{q}(t),t), \qquad i = 1,\dots,\ell \tag{3.11.7}$$

Observation. Note that Eqs. (3.11.6) and (3.11.7) are equations to which the local existence, uniqueness, and regularity theorems for differential equations can be immediately applied; this is not the case for Eq. (3.11.3), where the highest derivatives do not necessarily appear with constant coefficients: see also the final part of the proof of §3.8 p.196, to realize that this is really an inconvenience.

PROOF. We only discuss the case $U \subset \mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$, leaving the other two cases ($U \subset \mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or $U \subset \mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}, \ell_1 + \ell_2 = \ell$) as exercises. In any case, the proof is just an algebraic check. Equation (3.11.3) can be written by Eq. (3.11.2) as

$$\frac{d}{dt}p_i(t) = \frac{\partial \mathcal{L}}{\partial q_i}(\dot{\mathbf{q}}(t),\mathbf{q}(t),t), \qquad i = 1,\dots,\ell, \tag{3.11.8}$$

but, by Eq. (3.11.5), $\forall i = 1,\dots,\ell$,

$$\frac{\partial H}{\partial q_i} = \frac{\partial \mathcal{L}}{\partial q_i} - \sum_{j=1}^{\ell} \frac{\partial \mathcal{L}}{\partial \alpha_j}\frac{\partial \alpha_j}{\partial q_i} + \sum_{j=1}^{\ell} p_j \frac{\partial \alpha_j}{\partial q_i}, \tag{3.11.9}$$

and by Eqs. (3.11.2) and (3.11.4), implying $p_j \equiv \frac{\partial \mathcal{L}}{\partial \alpha_j}(\boldsymbol{\alpha}(\mathbf{p},\mathbf{q},t),\mathbf{q}(t),t)$, the two sums cancel and Eqs. (3.11.8) and (3.11.9) become Eq. (3.11.6).

Furthermore, by Eqs. (3.11.5) and (3.11.2),

$$\frac{\partial H}{\partial p_i} = \alpha_i + \sum_{j=1}^{\ell} p_j \frac{\partial \alpha_j}{\partial p_i} = \alpha_i = \dot{q}_i \tag{3.11.10}$$

i.e., Eq. (3.11.7) follows. mbe

The above proposition suggests a definition.

15 Definition. *Let V be an open set in $W \times \mathcal{R}^\ell \times \mathcal{R}$ (or $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$, $\ell_i > 0$)*[14] *and let H be a real-valued $C^\infty(V)$ function. H will be said to be a "regular Hamiltonian function" on V if the map Ψ transforming the point $(\boldsymbol{\pi}, \boldsymbol{\beta}, t) \in V$ into the point*

$$\boldsymbol{\psi}(\boldsymbol{\pi}, \boldsymbol{\beta}, t) = (\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \tag{3.11.11}$$

with

$$\alpha_i = \frac{\partial H}{\partial \pi_i}(\boldsymbol{\pi}, \boldsymbol{\beta}, t), \qquad i = 1, \ldots, \ell, \tag{3.11.12}$$

maps V in a neighborhood $U \subset \mathcal{R}^\ell \times \mathcal{R}$ (or $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$, $\ell_2 > 0$, respectively), $U = \Psi(V)$, invertibly and nonsingularly.[15]

If H is a regular Hamiltonian on V, the equations for the motion $t \to (\mathbf{p}(t), \mathbf{q}(t), t)$ in V:

$$\dot{p}_i(t) = -\frac{\partial H}{\partial \beta_i}(\mathbf{p}(t), \mathbf{q}(t), t) \tag{3.11.13}$$

$$\dot{q}_i(t) = \frac{\partial H}{\partial \pi_i}(\mathbf{p}(t), \mathbf{q}(t), t) \tag{3.11.14}$$

are called the "Hamiltonian differential equations" for the Hamiltonian H.

A proposition similar to Proposition 16 holds.

17 Proposition. *Let H be a regular Hamiltonian function on $V \subset RR^\ell \times \mathcal{R}$ (or $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$, $\ell_2 > 0$). Let $t \to (\mathbf{p}(t), \mathbf{q}(t), t)$ be a motion in V defined for $t \in [t_1, t_2]$ and verifying Eqs. (3.11.13) and (3.11.14). Setting*

$$(\boldsymbol{\pi}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t), \boldsymbol{\beta}, t) = \boldsymbol{\Psi}^{-1}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \tag{3.11.15}$$

for $(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \in U$ (see definition 15) and

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) = \sum_{j=1}^{\ell} \pi_j(\boldsymbol{\alpha}, \boldsymbol{\beta}, t)\alpha_j - H(\boldsymbol{\pi}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t), \boldsymbol{\beta}, t) \tag{3.11.16}$$

the motion in U, $t \to (\mathbf{a}(t), \mathbf{q}(t), t) = \Psi^{-1}(\mathbf{p}(t), \mathbf{q}(t), t)$, $t \in [t_1, t_2]$, verifies the equations:

$$\dot{\mathbf{q}}(t) = \mathbf{a}(t), \tag{3.11.17}$$

$$\frac{d}{dt}\Big(\frac{\partial \mathcal{L}}{\partial_{\alpha_i}}(\dot{\mathbf{q}}(t), \mathbf{q}(t), t)\Big) = \frac{\partial \mathcal{L}}{\partial_{\beta_i}}(\dot{\mathbf{q}}(t), \mathbf{q}(t), t), \quad i = 1, \ldots, \ell \tag{3.11.18}$$

---

[14] As in Definition 14, this definition makes sense without change if V is an open subset of $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$ (see Definition 14 and Observation (1) to Definition 14).

[15] i.e., with non vanishing Jacobian determinant. Such a Jacobian determinant is easily seen to be the determinant of the matrix $J_{ij} = (\partial^2 H / \partial \pi_i \partial \pi_j)$, $i, j = 1, \ldots, \ell$.

Proof. The proof is basically identical to that of Proposition 16.

Observations.
(1) Propositions 16 and 17 show that "a system of Lagrangian equations, regular on U, is equivalent to a system of Hamiltonian equations, regular on V, and vice versa". The sets U and V are related by the relations:
(i) V is the image of U via the map

$$\boldsymbol{\Xi} : (\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \to (\boldsymbol{\pi}, \boldsymbol{\beta}, t) = \big(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t), \boldsymbol{\beta}, t\big) \tag{3.11.19}$$

where $\mathcal{L}$ is the Lagrangian function on U.
(ii) U is the image of V via the map

$$\boldsymbol{\Psi} : (\boldsymbol{\pi}, \boldsymbol{\beta}, t) \to (\boldsymbol{\alpha}, \boldsymbol{\beta}, t) = \big(\frac{\partial H}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi}, \boldsymbol{\beta}, t), \boldsymbol{\beta}, t\big) \tag{3.11.20}$$

where H is the Hamiltonian function on V corresponding to $\mathcal{L}$.
(iii) $\mathcal{L}$ and the corresponding H are related by

$$H(\boldsymbol{\pi}, \boldsymbol{\beta}, t) = \sum_{i=1}^{\ell} \pi_i \alpha_i(\boldsymbol{\pi}, \boldsymbol{\beta}, t) - \mathcal{L}(\boldsymbol{\alpha}(\boldsymbol{\pi}, \boldsymbol{\beta}, t), \boldsymbol{\beta}, t), \tag{3.11.21}$$

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) = \sum_{i=1}^{\ell} \pi_i(\boldsymbol{\alpha}, \boldsymbol{\beta}, t)\alpha_i - H(\boldsymbol{\pi}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t), \boldsymbol{\beta}, t). \tag{3.11.22}$$

(2) In the applications met so far, $\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t)$ has always had the form

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) = \frac{1}{2} \sum_{i,j=1}^{\ell} g_{ij}(\boldsymbol{\beta})\alpha_i\alpha_j - V(\boldsymbol{\beta}) \tag{3.11.23}$$

where g is a positive definite matrix. $\mathcal{L}$ is usually defined in neighborhoods $U = \mathcal{R} \times U_0 \times \mathcal{R}$, $U_0 \subset \mathcal{T}^\ell$ (or $U_0 \subset \mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}$, $\ell_1 + \ell_2 = \ell$, $\ell_i > 0$) open. Eq. (3.11.23) is regular on U because Eq. (3.11.19) [or Eq. (3.11.2)] becomes

$$\pi_i = \sum_{j=1}^{\ell} g_{ij}(\boldsymbol{\beta})\alpha_j, \qquad i = 1, \ldots, 1, \tag{3.11.24}$$

which is invertible and nonsingular if thought of as defining [see Eq. (3.11.19)] a map of U onto $V = \mathcal{R}^\ell \times U_0 \times \mathcal{R}$: this is so by virtue of Proposition 11, §3.7, p.182, on the kinetic matrices (implying $\det g(\boldsymbol{\beta}) \neq 0$).
The Hamiltonian function associated with Eq. (3.11.23) is, by Eqs. (3.11.24) and (3.11.21),

$$H(\boldsymbol{\pi}, \boldsymbol{\beta}, t) = \frac{1}{2} \sum_{i,j=1}^{\ell} (g(\boldsymbol{\beta})^{-1})_{ij}\, \pi_i\, \pi_j + V(\boldsymbol{\beta}), \tag{3.11.25}$$

where $g(\boldsymbol{\beta})^{-1}$ is the inverse matrix to $g(\boldsymbol{\beta})$.
(3) In the case of the Lagrangian (3.11.23), Eq. (3.11.2) [i.e., Eq. (3.11.24)] is simply the condition expressing that the gradient of the function of $\boldsymbol{\alpha} \in \mathcal{R}^\ell$,

$$\boldsymbol{\alpha} \to \boldsymbol{\pi} \cdot \boldsymbol{\alpha} - \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) \tag{3.11.26}$$

vanishes. One can check that for such a value of $\boldsymbol{\alpha}$, Eq. (3.11.26) actually reaches its only absolute maximum [Note that in the case considered here, Eq. (3.11.26) is a quadratic form in $\boldsymbol{\alpha}$ plus a linear form in $\boldsymbol{\alpha}$.] So

$$H(\boldsymbol{\pi}, \boldsymbol{\beta}, t) = \max_{\boldsymbol{\alpha} \in \mathcal{R}^\ell} (\boldsymbol{\pi} \cdot \boldsymbol{\alpha} - \mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t)) \tag{3.11.27}$$

when $\mathcal{L}$ is given by Eq. (3.11.23) or, more generally, whenever the function of Eq. (3.11.26) has only one stationarity point in a which is a maximum (exercise). Similarly,

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}, t) = \max_{\boldsymbol{\pi} \in \mathcal{R}^\ell} (\boldsymbol{\pi} \cdot \boldsymbol{\alpha} - H(\boldsymbol{\pi}, \boldsymbol{\beta}, t)) \tag{3.11.28}$$

if H is given by Eq. (3.11.25) or, more generally, whenever the function of $\boldsymbol{\pi}$ inside the parenthesis on the right-hand side has only one stationarity point in a which is a maximum.
Equations (3.11.27) and (3.11.28) are often called "Legendre's duality" or "Legendre's transformations" on $\mathcal{L}$ or H, respectively.
(4) Definitions 14 and 15 and Propositions 16 and 17 assume a simpler form if one is interested in Lagrangian or Hamiltonian functions not explicitly depending on time and defined on sets U or V of the form $\widehat{U} \times J$ or $\widehat{V} \times J$ with $J = \{\text{open interval in } \mathcal{R}\}$ and $\widehat{U}, \widehat{V} \subset \mathcal{R}^\ell$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}), \ell_1 + \ell_2 = \ell$, open sets.
In such cases, the t parameter can be eliminated from the definition of the sets U, V (replacing them by $\widehat{U}, \widehat{V}$) and of the maps $\boldsymbol{\Xi}, \boldsymbol{\Psi}$ in Definitions 14 and 15, and $\mathcal{L}$ or H will be functions in $C^\infty(\widehat{U})$ or on $C^\infty(\widehat{V})$. $\mathcal{L}$ or H will be called "time-independent" Lagrangian or Hamiltonian functions and they generate autonomous Lagrangian or Hamiltonian equations via Eqs. (3.11.3), (3.11.6), and (3.11.7).
When $\widehat{V} = \mathcal{R}^\ell \times U_0$, the space $\widehat{V}$ is usually called the "phase space" if it is regarded as the initial data space for some time-independent Hamiltonian equations: this name is often used even when V is just an open set (not necessarily of the form $\mathcal{R}^\ell \times U_0$). Similarly, when $\widehat{U} = \mathcal{R}^\ell \times U_0$, the space $\widehat{U}$ is called the "data space" if it is regarded as the initial data space for a time-independent Lagrangian equation.

The formal wording of the above concepts is straightforward and will be left to the reader. We shall freely refer to time-independent Lagrangian or Hamiltonian functions and equations on the data space or the phase space.

It is interesting to note the following abstract version of the energy conservation theorem.

18 proposition. *Consider a system of Hamiltonian equations in a neighborhood* $U = \mathcal{R} \times U_0 \times \mathcal{R}$, $U_0 \subset \mathcal{R}^\ell$, *(or* $U_0 \subset \mathcal{T}^\ell$ *or* $U_0 \subset \mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}$, $\ell_1 + \ell_2 = \ell$*) and let* $(\boldsymbol{\pi}, \boldsymbol{\beta}, t) \to H(\boldsymbol{\pi}, \boldsymbol{\beta}, t)$ *be the (regular) Hamiltonian function. If* $t \to (\mathbf{p}(t), \mathbf{q}(t), t) \in U$, $t \in [t_1, t_2]$, *is a motion verifying in* U *the Hamiltonian equations, then*

$$\frac{d}{dt}(H(\mathbf{p}(t), \mathbf{q}(t), t)) = \frac{\partial H}{\partial t}(\mathbf{p}(t), \mathbf{q}(t), t), \tag{3.11.29}$$

Hence, if H *is time independent, i.e.,* $H(\boldsymbol{\pi}, \boldsymbol{\beta}, t) \equiv h(\boldsymbol{\pi}, \boldsymbol{\beta})$ *for some* $h \in C^\infty(\mathcal{R}^\ell \times U_0)$*, Eq. (3.11.29) implies the existence of a constant* E*, depending on the motion under investigation, such that*

$$h(\mathbf{p}(t), \mathbf{q}(t)) = E, \qquad t \in [t_1, t_2]. \tag{3.11.30}$$

Observations.
(1) In the cases met so far, the Lagrange function had the form of Eq. (3.11.23) and $\frac{1}{2}\sum_{i,j=1}^{\ell} g_{ij}(\mathbf{q}(t))\dot{q}_i(t)\dot{q}_j(t)$ had the interpretation of kinetic energy $T(t)$ of the motion, while $V(\mathbf{q}(t))$ had the interpretation of potential energy $V(t)$. Furthermore, the relation between $\mathbf{p}(t)$ and $\dot{\mathbf{q}}(t)$ was [Eq. (3.11.24)]:

$$\mathbf{p}(t) = g(\mathbf{q}(t))\dot{\mathbf{q}}(t). \tag{3.11.31}$$

Then, by Eq. (3.11.31),

$$\frac{1}{2}\sum_{i,j=1}^{\ell} \left(g(\mathbf{q}(t))^{-1}\right)_{ij} p_i(t)p_j(t) = \frac{1}{2}\sum_{i,j=1}^{\ell} g(\mathbf{q}(t))_{ij}\dot{q}_i(t)\dot{q}_j(t) \equiv T(t). \tag{3.11.32}$$

Hence Eq. (3.11.30) becomes

$$T(t) + V(t) = E. \tag{3.11.33}$$

(2) When a system of N points without constraints, with the Lagrangian function

$$\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{2}\sum_{i=1}^{\ell} m_i \alpha_i^2 - V(\boldsymbol{\beta}) \tag{3.11.34}$$

is considered, we see that $\boldsymbol{\pi} = (\boldsymbol{\pi}^{(1)}, \ldots, \boldsymbol{\pi}^{(N)})$ with $\boldsymbol{\pi}^{(i)} = m_i \boldsymbol{\alpha}^{(i)}, i = 1, \ldots, N$, so that if $t \to \mathbf{x}(t)$, $t \in [t_1, t_2]$, is a system motion:

$$p_i(t) = m_i\, \dot{q}_i(t), \qquad i = 1, \ldots, \ell \tag{3.11.35}$$

which explains the name "generalized momenta" given to the variables $\boldsymbol{\pi}_i$ in general.

The variables π_i are also called the "conjugated momenta" with respect to β_i, $i = 1, \ldots, \ell$, and the 2ℓ variables $(\boldsymbol{\pi}, \boldsymbol{\beta})$ are called "canonical" variables in the phase space of a Hamiltonian equation.
The word conjugation is used here because of the obvious symmetric role played by the p and q variables in the Hamiltonian equations. This symmetry could be used to build even more abstract structures associated with the theory of mechanical equations of motion for conservative systems; however, this aim will not be pursued here.

Proof. In fact,

$$\begin{aligned}\frac{d}{dt}H(\mathbf{p}(t),\mathbf{q}(t),t) &= \frac{\partial H}{\partial t}(\mathbf{p}(t),\mathbf{q}(t),t) + \sum_{i=1}^{\ell}\big(\frac{\partial H}{\partial \pi_i}\,\dot{p}_i + \frac{\partial H}{\partial \beta_i}\,\dot{q}_i\big) \\ &= \frac{\partial H}{\partial t}(\mathbf{p}(t),\mathbf{q}(t),t) + \sum_{i=1}^{\ell}(\dot{q}_i\dot{p}_i - \dot{p}_i\dot{q}_i\big) = \frac{\partial H}{\partial t}(\mathbf{p}(t),\mathbf{q}(t),t)\end{aligned} \tag{3.11.36}$$

mbe

Another consequence, already mentioned in Problem 10, §2.24, p.137, of the symmetry of Hamiltonian equations is the following:

19 Proposition. *Let $V = \mathcal{R}^\ell \times U_0$ with U_0 open subset of $\mathcal{R}^\ell$ (or $\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}, \ell_1 + \ell_2 = \ell$). Let $h \in C^\infty(\mathcal{R}^\ell \times U_0)$ be a time-independent regular Hamiltonian function.*[16] *Call $S_t(\boldsymbol{\pi}, \boldsymbol{\beta})$ the point into which the initial datum $(\boldsymbol{\pi}, \boldsymbol{\beta})$ evolves through the equations:*

$$\dot{\mathbf{p}} = -\frac{\partial h}{\partial \boldsymbol{\beta}}(\mathbf{p}, \mathbf{q}), \qquad \dot{\mathbf{q}} = \frac{\partial h}{\partial \boldsymbol{\pi}}(\mathbf{p}, \mathbf{q}). \tag{3.11.37}$$

Suppose that for $\tau \in [0, t]$, the data $(\boldsymbol{\pi}, \boldsymbol{\beta}) \in A \subset U$ are such that $S_\tau(\boldsymbol{\pi}, \boldsymbol{\beta}) \in U$, i.e., $S_\tau A \subset U$ if $\tau \in [0, t]$, i.e. the evolution of the points in A takes place inside U for all $\tau \in [0, t]$, and suppose that A is measurable; then

$$\text{volume}\, S_t A = \int_{S_t A} d\mathbf{p} d\mathbf{q} = \text{volume}\, A. \tag{3.11.38}$$

Observation. This is read by saying "the Hamiltonian flow preserves the phase space volume" and it is called the "Liouville theorem".

Proof. This is a consequence of the fact that the Hamiltonian equations have zero divergence: $\sum_{i=1}^{\ell} -\frac{\partial^2}{\partial \pi_i \partial \beta_i} + \sum_{i=1}^{\ell} \frac{\partial^2}{\partial \beta_i \partial \pi_i} = 0$ (see the hint to Problem 10, §2.24, where the argument is given in detail).

mbe

[16] see observation (4) to Proposition 17; the function $H(\boldsymbol{\pi}, \boldsymbol{\beta}, t) \equiv h(\boldsymbol{\pi}, \boldsymbol{\beta})$ is a regular Hamiltonian on $\cong V \times \mathcal{R}$ in the sense of the Definition 15, p.214.

A corollary to the above proposition is the following.

20 Proposition. *Given the same assumptions as in Proposition 19, suppose, also, that the set of the* $(\boldsymbol{\pi}, \boldsymbol{\beta})$ *such that* $h(\boldsymbol{\pi}, \boldsymbol{\beta}) < E$ *is a set* Ω_E *whose closure in* $\mathcal{R}^\ell \times \mathcal{R}^\ell$ *(or* $\mathcal{R}^\ell \times \mathcal{T}^\ell$ *or* $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}'^{\ell_2}), \ell_1 + \ell_2 = \ell$*) is contained in* V *and is bounded. Then given any* $(\boldsymbol{\pi}_0, \boldsymbol{\beta}_0) \in \Omega_E$, $t_0 > 0$ *and a neighborhood* $W \subset \Omega_E$ *of* $(\boldsymbol{\pi}_0, \boldsymbol{\beta}_0)$*, there exists* $t > t_0$ *such that* $S_t W \cap W \neq \emptyset$.

Observation. So "if the energy E surface is bounded", close to every point "inside" it there is another point coming "as close" after any given time: this is Poincaré's recursion theorem.
If the system contains N (e.g. $\sim 10^{24}$) points enclosed in a box (modeled by a potential tending very quickly to $+\infty$ outside the box) and if it is initially in a configuration $\boldsymbol{\xi}$ in which all points are confined to the left half of the box (say), then as close to it as we wish there is *another* configuration which evolves so that, waiting "long enough", we shall be surprised to see that all the particles will again occupy the left half of the box. This nice paradox ("Zermelo's paradox") gave some problems to Boltzmann.

Proof. The proof is a very simple consequence of Proposition 19 and is described in greater generality (for divergenceless differential equations) in Problem 11, §2.24, p.138 (see hint). mbe

In connection with the Hamiltonian equations, the notion of "canonical transformation" plays an important role. A transformation of coordinates is canonical when it leaves the structure of the Hamiltonian equations unchanged. Such a notion has remarkable importance in the algorithms used in the theory of perturbations, which we shall introduce in Chapter 5.

16 Definition. *Let* V *be an open set in* $\mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$ *(or in* $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ *or* $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$*) and let* H *be a regular Hamiltonian function on* V *(see Definition 15).*
Suppose that on V *a* C^∞ *map* $\mathcal{C}$ *is defined such that:*
(i) The image of $(\mathbf{p}, \mathbf{q}, t) \in V$ *has the form* $(\boldsymbol{\pi}, \boldsymbol{\kappa}, t) = \mathcal{C}(\mathbf{p}, \mathbf{q}, t)$*, i.e.,* $\mathcal{C}$ *is an "isochronous map" (since it does not affect* t*).*
(ii) The map $\mathcal{C}$ *maps* V *onto* $W = \mathcal{C}(V)$*, which is an open subset of* $\mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$ *(or in* $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ *or* $\mathcal{R}^\ell \times (\mathcal{R}^{\ell'_1} \times \mathcal{T}^{\ell'_2}) \times \mathcal{R}$, $\ell'_1 + \ell'_2 = \ell$*) and it is invertible and nonsingular,* [17] *i.e.,* $\mathcal{C}$ *is a regular change of coordinates on* V.
(iii) There is a real-valued function $H' \in C^\infty(W)$ *such that if* $t \to (\mathbf{p}(t), \mathbf{q}(t), t) \in V, t \in [t_1, t_2]$*, is any motion in* V *verifying the Hamiltonian equations with Hamiltonian* H*, then* $t \to (\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t), t) = \mathcal{C}(\mathbf{p}(t), \mathbf{q}(t), t) \in W$, $t \in [t_1, t_2]$*, verifies the Hamiltonian equations relative to* H' *and vice versa. One says that* $\mathcal{C}$ *is a "canonical transformation of* V *in* W *with respect to the pair of conjugate Hamiltonians* H *and* H'*".*

[17] i.e., its Jacobian determinant does not vanish. Hence, $\mathcal{C}^{-1}$ has the same properties by the implicit function theorem.

Observation. In general, if a map $\mathcal{C}$ is canonical for the pair H, H', it will not be canonical for the pair (H, H'') no matter how H'' is chosen, if $H'' \neq H'$ (for an example, see Problem 38, at the end of this section).

It is therefore tempting to call "completely canonical" a map $\mathcal{C}$ between V and W such that *for any choice* of a Hamiltonian function H on V, one can find a conjugated Hamiltonian function H' on W in some standard way (Levi-Civita).

We shall make the notion of "complete canonicity" precise only in the simple case of "time-independent" canonical transformations.

17 Definition. *Let $V = \widehat{V} \times R$ be an open subset of $\mathcal{R}^\ell \times \mathcal{R}^\ell \times R$ (or of $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or of $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}^\ell$, $\ell_1 + \ell_2 = \ell$) and let $\mathcal{C}$ have the form $\mathcal{C}(\mathbf{p}, \mathbf{q}, t) = (\widehat{\mathcal{C}}(\mathbf{p}, \mathbf{q}), t)$ with $\widehat{\mathcal{C}}$ being a regular change of coordinates between $\widehat{V}$ and its image $\widehat{W}$ $\widehat{W} \subset \mathcal{R}^\ell \times \mathcal{R}^\ell$ (or $\widehat{W} \subset \mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\widehat{W} \subset \mathcal{R}^\ell \times (\mathcal{R}^{\ell_1'} \times \mathcal{T}^{\ell_2'}), \ell_1' + \ell_2' = \ell$).*
We shall say that $\widehat{\mathcal{C}}$ is a "completely canonical time-independent" or, simply, a "completely canonical" transformation if $\widehat{\mathcal{C}}$ is a transformation which conjugates canonically every regular Hamiltonian function H on V with

$$H'(\boldsymbol{\pi}, \boldsymbol{\kappa}, t) \stackrel{def}{=} H(\widehat{\mathcal{C}}^{-1}(\boldsymbol{\pi}, \boldsymbol{\kappa}), t), \qquad \forall (\boldsymbol{\pi}, \boldsymbol{\kappa}) \in \widehat{W}. \tag{3.11.39}$$

Observation. In other words, a time-independent completely canonical transformation is one with the property that any Hamiltonian function is conjugated to itself computed in the new coordinates.

The following proposition provides a very general method of construction of canonical transformations and of completely canonical transformations.

21 Proposition. *Let H be a regular Hamiltonian function on the open set $V \subset \mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$ (or in $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{R}^{\ell_2}) \times \mathcal{R}, \ell_1 + \ell_2 = \ell$). Let $F \in C^\infty(\mathcal{R}^{2\ell+1})$ be a function denoted*

$$(\mathbf{q}, \boldsymbol{\kappa}, t) \to F(\mathbf{q}, \boldsymbol{\kappa}, t) \in \mathcal{R} \tag{3.11.40}$$

For $i = 1, \ldots, \ell$, set $t' = t$ and

$$p_i = \frac{\partial F}{\partial q_i}(\mathbf{q}, \boldsymbol{\kappa}, t), \qquad \pi_i = -\frac{\partial F}{\partial \kappa_i}(\mathbf{q}, \boldsymbol{\kappa}, t) \tag{3.11.41}$$

and assume that Eq. (3.11.41) establishes a one-to-one map $\mathcal{C}_F$ between $(\mathbf{p}, \mathbf{q}, t) \in V$ and $(\boldsymbol{\pi}, \boldsymbol{\kappa}, t) = \mathcal{C}_F(\mathbf{p}, \mathbf{q}, t) \in W$. Suppose that $\mathcal{C}_F$ is a regular change of coordinates[18] *between V and W ($W \subset \mathcal{R}^\ell \times \mathcal{R}^\ell \times \mathcal{R}$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{R}^{\ell_2}) \times \mathcal{R}, \ell_1 + \ell_2 = \ell$). Then if we define $(\mathbf{p}(\boldsymbol{\pi}, \boldsymbol{\kappa}, t), q(\boldsymbol{\pi}, \boldsymbol{\kappa}, t)) \equiv \mathcal{C}_F^{-1}(\boldsymbol{\pi}, \boldsymbol{\kappa}, t)$ and*

[18] i.e., it is one-to-one and with non vanishing Jacobian determinant.

$$H'(\boldsymbol{\pi},\boldsymbol{\kappa},t) = H(\mathbf{p}(\boldsymbol{\pi},\boldsymbol{\kappa},t),\mathbf{q}(\boldsymbol{\pi},\boldsymbol{\kappa},t),t) + \frac{\partial F}{\partial t}(\mathbf{q}(\boldsymbol{\pi},\boldsymbol{\kappa},t),\boldsymbol{\kappa},t), \qquad (3.11.42)$$

the map $\mathcal{C}_F$ is a canonical transformation of V onto W with respect to H and H'.

Observations.
(1) Note that F is required to be in $C^\infty(\mathcal{R}^{2\ell+1})$ even when V is in $\mathcal{R}^\ell \times \mathcal{T}^\ell \times \mathcal{R}$ or in $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}) \times \mathcal{R}$, $\ell_1 + \ell_2 = \ell$. Recall that the points on a torus are, in the present contexts, always thought of as described in "flat or angular coordinates", i.e., by thinking of the torus $\mathcal{T}^\ell$ as obtained by identifying mod 2π the points of $\mathcal{R}^\ell$ (see Definition 14, p.211).
(2) From the proof of Proposition 21, it will follow that other coordinate transformations analogous to Eq. (3.11.41) are canonical: for instance, from a function $\Phi \in C^\infty(\mathcal{R}^{2\ell+1})$,

$$(\mathbf{q},\boldsymbol{\pi},t) \to \Phi(\mathbf{q},\boldsymbol{\pi},t) \in \mathcal{R}, \qquad (3.11.43)$$

one builds a canonical transformation[19] $\mathcal{C}_\Phi$ by setting, $\forall i = 1,2,\ldots,\ell$,

$$p_i = \frac{\partial \Phi}{\partial q_i}(\mathbf{q},\boldsymbol{\pi},t), \qquad \kappa_i = \frac{\partial \Phi}{\partial \pi_i}(\mathbf{q},\boldsymbol{\pi},t), \qquad (3.11.44)$$

$$H'(\boldsymbol{\pi},\boldsymbol{\kappa},t) = H(\mathbf{p}(\boldsymbol{\pi},\boldsymbol{\kappa},t),\mathbf{q}(\boldsymbol{\pi},\boldsymbol{\kappa},t),t) + \frac{\partial \Phi}{\partial t}(\mathbf{q}(\boldsymbol{\pi},\boldsymbol{\kappa},t),\boldsymbol{\pi},t) \qquad (3.11.45)$$

where we denote $(\mathbf{p}(\boldsymbol{\pi},\boldsymbol{\kappa},t),\mathbf{q}(\boldsymbol{\pi},\boldsymbol{\kappa},t),t) = \mathcal{C}_\Phi(\boldsymbol{\pi},\boldsymbol{\kappa},t)$
Similarly, with analogous notations, if $\Psi \in C^\infty(\mathcal{R}^{2\ell+1})$, setting

$$(\mathbf{p},\boldsymbol{\kappa},t) \to \Psi(\mathbf{p},\boldsymbol{\kappa},t) \in \mathcal{R}, \qquad (3.11.46)$$

one defines a canonical transformation[19] $\mathcal{C}_\Psi$ by setting $\forall i = 1,\ldots,\ell$:

$$\begin{aligned} q_i &= -\frac{\partial \Psi}{\partial p_i}(\mathbf{p},\boldsymbol{\kappa},t), \qquad \pi_i = -\frac{\partial \Psi}{\partial \kappa_i}(\mathbf{p},\boldsymbol{\kappa},t), \\ H' &= H + \frac{\partial \Psi}{\partial t}, \end{aligned} \qquad (3.11.47)$$

and if $R \in C^\infty(\mathcal{R}^{2\ell+1})$,

$$(\mathbf{p},\boldsymbol{\pi},t) \to R(\mathbf{p},\boldsymbol{\pi},t) \in \mathcal{R}, \qquad (3.11.48)$$

defines a canonical transformation[19] $\mathcal{C}_R$ by setting, $\forall i = 1,2,\ldots,\ell$,

[19] between regions where the regularity, invertibility and nonsingularity requirements for the maps $\mathcal{C}_\Phi$ (or, see below, $\mathcal{C}_\Psi, \mathcal{C}_R$) similar to those put on $\mathcal{C}_F$ are verified. Such regions V, W may be very small or even nonexistent: in the last cases no canonical transformation is really associated with F, Φ, Ψ, R.

$$\begin{aligned} q_i &= -\frac{\partial R}{\partial p_i}(\mathbf{p}, \boldsymbol{\pi}, t), \qquad \kappa_i = \frac{\partial R}{\partial \pi_i}(\mathbf{p}, \boldsymbol{\pi}, t), \\ H' &= H + \frac{\partial R}{\partial t}. \end{aligned} \tag{3.11.49}$$

(3) However, it will appear that the class of canonical transformations built starting from F as described by Proposition 20 is not essentially less ample than that obtained by adding to it the canonical transformations associated with the functions Φ, Ψ and R as described in the preceding observation. With some natural exceptions, to every F it is possible to associate a Φ, a Ψ, and an R producing the same canonical transformation.
(4) If F is time independent, then $\mathcal{C}_F$ defines a completely canonical (time-independent) map.
(5) F is in general called a "generating function" of $\mathcal{C}_F$. So one calls also the functions Φ, Ψ, R above.

PROOF. The proof by direct check is of course possible. However, if performed straightforwardly, it quickly becomes quite intricate. It is certainly more convenient to proceed in the following elegant fashion, which also exhibits a new form of the least-action principle: the "Hamilton's principle".

Let $\mathcal{M}^V = \mathcal{M}_{t_1 t_2}(\mathbf{p}_1, \mathbf{q}_1, t; \mathbf{p}_2, \mathbf{q}_2, t_2; V) = \{$set of the motions in V having the form $t \to \mathbf{m}(t) = (\mathbf{p}(t), \mathbf{q}(t), t) \in V$, $t \in [t_1, t_2]$ and such that $\mathbf{p}(t_1) = \mathbf{p}_1, \mathbf{q}(t_1) = \mathbf{q}_1, \mathbf{p}(t_2) = \mathbf{p}_2, \mathbf{q}(t_2) = \mathbf{q}_2\}$ ("synchronous motions $\mathbf{m}$ in Vs"). Consider the function on $\mathcal{M}^V$:

$$\mathcal{S}(\mathbf{m}) = \int_{t_1}^{t_2} \big(\sum_{i=1}^{\ell} p_i(t), \dot{q}_i(t) - H(\mathbf{p}(t), \mathbf{q}(t), t)\big)\, dt. \tag{3.11.50}$$

With the methods of §2.2.1 and §3.4 by now familiar, one checks that the stationarity condition for $\mathcal{S}$ on $\mathbf{m}$ in $\mathcal{M}^V$ is simply that the motion $\mathbf{m}$ verifies the Hamiltonian equations in V with Hamiltonian function H [which are, essentially, the Euler-Lagrange equations for the action of Eq. (3.11.50)].

Now let $t \to \boldsymbol{\mu}(t) = (\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t), t) = \mathcal{C}_F(\mathbf{p}(t), \mathbf{q}(t), t)$, $t \in [t_1, t_2]$ be the image motion of a motion $\mathbf{m} \in \mathcal{M}^V$: it is a motion in

$\mathcal{C}_F(\mathbf{p}_1, \mathbf{q}_1, t_1)\mathcal{M}^W = \mathcal{C}_F(\mathcal{M}^V) = \mathcal{M}_{t_1 t_2}(\mathcal{C}_F(\mathbf{p}_1, \mathbf{q}_1, t_1), \mathcal{C}_F(\mathbf{p}_2, \mathbf{q}_2, t_2); W)$

If $\boldsymbol{\mu}$ verifies the Hamiltonian equations for some Hamiltonian H' on W in $\mathbf{m} \in \mathcal{M}^W$, it must make the action

$$\Sigma(\boldsymbol{\mu}) = \int_{t_1}^{t_2} \big\{\sum_{i=1}^{\ell} \pi_i(t)\dot{\kappa}_i(t) - H'(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t), t)\big\}\, dt \tag{3.11.51}$$

stationary. A *sufficient* condition for this to occur is that

$$\mathcal{S}(\mathbf{m}) = \Sigma(\mathcal{C}_F(\mathbf{m})) + \text{constant}, \qquad \forall \mathbf{m} \in \mathcal{M}^V, \tag{3.11.52}$$

of course. Equation (3.11.52) is certainly verified if the differential form on V:

$$\sum_{i=1}^{\ell} p_i dq_i - H(\mathbf{p}, \mathbf{q}, t)\, dt \tag{3.11.53}$$

and the differential form

$$\sum_{i=1}^{\ell} \pi_i d\kappa_i - H'(\boldsymbol{\pi}, \boldsymbol{\kappa}, t) dt \tag{3.11.54}$$

are transformed into each other by the transformation $\mathcal{C}_F$, *up to a total differential.* This condition can be imposed by requiring the existence of a function G on W such that

$$\sum_{i=1}^{\ell} p_i dq_i - H(\mathbf{p}, \mathbf{q}, t)\, dt = \sum_{i=1}^{\ell} \pi_i d\kappa_i - H'(\boldsymbol{\pi}, \boldsymbol{\kappa}, t) dt + dG \tag{3.11.55}$$

where $(\mathbf{p}, \mathbf{q}, t)$ are to be thought of as functions of $(\boldsymbol{\pi}, \boldsymbol{\kappa}, t)$ via the transformation $\mathcal{C}_F$.

To use Eq. (3.11.55), it is more convenient to think of G as a function of $(\mathbf{q}, \boldsymbol{\kappa}, t)$ instead of $(\boldsymbol{\pi}, \boldsymbol{\kappa}, t)$ via Eq. (3.11.41); i.e., set $\widetilde{G}(\mathbf{q}, \boldsymbol{\kappa}, t) = G(\frac{\partial F}{\partial \mathbf{q}}(\mathbf{q}, \boldsymbol{\kappa}, t), \boldsymbol{\kappa}, t)$. Then it follows from Eq. (3.11.55) that

$$d\widetilde{G} = \sum_{i=1}^{\ell} p_i dq_i - \sum_{i=1}^{\ell} \pi_i d\kappa_i - (H - H')dt, \tag{3.11.56}$$

so we realize that Eq. (3.11.56) holds if and only if there is a function $\widetilde{G}$ which is such that, $\forall i = 1, \dots, \ell$,

$$p_i = \frac{\partial \widetilde{G}}{\partial q_i}, \qquad \pi_i = -\frac{\partial \widetilde{G}}{\partial \kappa_i}, \qquad H - H' = \frac{\partial \widetilde{G}}{\partial t} \tag{3.11.57}$$

thinking the coefficients of the right-hand side differentials in Eq. (3.11.56) as functions of $\mathbf{q}, \boldsymbol{\kappa}, t$, via Eq. (3.11.41). Such relations are satisfied by the function F, setting $\widetilde{G} = F$. mbe

Observations. Subtracting the differential $d(\sum_{i=1}^{\ell} p_i q_i)$ from both sides of Eq. (3.11.56) and thinking of

$$\Psi = F - \sum_{i=1}^{\ell} p_i q_i \tag{3.11.58}$$

as a function of $\mathbf{p}, \boldsymbol{\kappa}, t$ via Eq. (3.11.41)[20] one finds that the transformation $\mathcal{C}_F$ may also be thought of as $\mathcal{C}_\Psi$ described by Eq. (3.11.47). Similarly, setting

[20] assuming that the necessary inversions can actually be made.

$$\Phi = F + \sum_{i=1}^{\ell} \pi_i d\kappa_i \tag{3.11.59}$$

and thinking of Φ as a function of $(\mathbf{q}, \boldsymbol{\pi}, t)$ via Eq. (3.11.41)[20] one finds that $\mathcal{C}_F$ may also be thought of as $\mathcal{C}_\Phi$ described by Eq. (3.11.44). Finally, setting

$$R = F + \sum_{i=1}^{\ell} \pi_i d\kappa_i - \sum_{i=1}^{\ell} p_i q_i \tag{3.11.60}$$

and thinking of R as a function of $(\mathbf{p}, \boldsymbol{\pi}, t)$ via Eq. (3.11.41),[20] one finds that the $\mathcal{C}_F$ may also be thought of as $\mathcal{C}_R$ described by Eq. (3.11.49).

In the problems at the end of this section, it will appear that the inversions mentioned above (see footnote 20) can be performed at least in small regions under the respective conditions that the matrices $\frac{\partial^2 F}{\partial \kappa_i \partial \kappa_j}$, $\frac{\partial^2 \Phi}{\partial \kappa_i \partial q_j}$, $\frac{\partial^2 \Psi}{\partial q_i \partial q_j}$ have non vanishing determinants.

This somewhat clarifies Observation (3) to Proposition 21. A complete clarification arises from the analysis of Problems (6)-(11) at the end of this section. The reader should try to think of these observations again after looking at the problems.

A simple corollary to the proof of Proposition 21 is the following.

22 Proposition. *Let $(\boldsymbol{\pi}, \boldsymbol{\kappa}) \to \mathcal{C}(\boldsymbol{\pi}, \boldsymbol{\kappa})$ be a nonsingular invertible C^∞ map of the open set $V \subset \mathcal{R}^{2\ell}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2})$, $\ell_1 + \ell_2 = \ell$, onto $W \subset \mathcal{R}^{2\ell}$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell'_1} \times \mathcal{T}^{\ell'_2})$, $\ell'_1 + \ell'_2 = \ell$. Write $\mathcal{C}$ explicitly as*

$$\mathbf{p} = \mathbf{P}(\pi, \boldsymbol{\kappa}), \qquad \mathbf{q} = \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa}) \tag{3.11.61}$$

and consider the differential form on V,

$$\boldsymbol{\pi} \cdot d\boldsymbol{\kappa} - \mathbf{p} \cdot d\mathbf{q} \equiv \sum_{i=1}^{\ell} (\pi_i d\kappa_i - p_i dq_i) \tag{3.11.62}$$

Write it as $-\sum_i (X_i d\pi_i + Y_i d\kappa_i)$ with

$$X_j = \mathbf{P}(\boldsymbol{\pi}, \boldsymbol{\kappa}) \cdot \frac{\partial \mathbf{Q}}{\partial \pi_i}(\boldsymbol{\pi}, \boldsymbol{\kappa}), \qquad Y_i = \mathbf{P}(\boldsymbol{\pi}, \boldsymbol{\kappa}) \cdot \frac{\partial \mathbf{Q}}{\partial \kappa_i}(\boldsymbol{\pi}, \boldsymbol{\kappa}) - \pi_i. \tag{3.11.63}$$

Suppose that the form in Eq. (3.11.62) is exact: i.e. $\forall\, i, j = 1, \ldots, \ell$

$$\frac{\partial X_i}{\partial \pi_j} = \frac{\partial X_j}{\partial \pi_i}, \quad \frac{\partial X_i}{\partial \kappa_j} = \frac{\partial Y_j}{\partial \pi_i}, \quad \frac{\partial Y_i}{\partial \kappa_j} = \frac{\partial Y_j}{\partial \kappa_i}, \tag{3.11.64}$$

Then $\mathcal{C}$ is a completely canonical time-independent map.
In particular, if $\boldsymbol{\pi} \cdot d\boldsymbol{\kappa} - \mathbf{p} \cdot d\mathbf{q} = 0$ the map $\mathcal{C}$ is completely canonical: it is called "homogeneous" in the variables $(\boldsymbol{\kappa}, \mathbf{q})$.
Similar results hold if $\mathbf{p}\, d\mathbf{q} + \boldsymbol{\kappa} \cdot f\boldsymbol{\pi}$, or $-\mathbf{q} \cdot d\mathbf{p} + \boldsymbol{\kappa} \cdot d\boldsymbol{\pi}$, or $-\mathbf{q} \cdot d\mathbf{p} - \boldsymbol{\pi} \cdot d\boldsymbol{\kappa}$

are exact differentials: one similarly defined the homogeneous canonical maps with respect to $(\mathbf{q},\boldsymbol{\pi})$*, or* $(\mathbf{p},\boldsymbol{\pi})$*, or* $(\mathbf{p},\boldsymbol{\kappa})$*.*

Observations.
(1) If $\mathcal{C}$ is as above and homogeneous in $(\boldsymbol{\kappa},\mathbf{q})$ variables, then it cannot generated by a generating function $F(\boldsymbol{\kappa},\mathbf{q})$ as in Eq. (3.11.41). The vanishing of the differential in Eq. (3.11.62) and the Eqs. (3.11.41), (3.11.42) written as

$$dF = \boldsymbol{\pi}\cdot d\boldsymbol{\kappa} - \mathbf{p}\cdot d\mathbf{q} + (H'-H)dt \tag{3.11.65}$$

imply that $dF=(H'-H)dt$, i.e. $H'-H$ is a function of t only and so is F as well, so that Eq. (3.11.41) gives $\boldsymbol{\pi}=\mathbf{0}$, $\mathbf{p}=\mathbf{0}$ which is obviously not usable to define an invertible map between $\mathbf{q},\mathbf{p}$ and $\boldsymbol{\kappa},\boldsymbol{\pi}$.
(2) If $\mathcal{C}$ is homogeneous as in Observation (1), it might be generated by functions $\Phi(\boldsymbol{\pi},\mathbf{q})$ or $\Psi(\boldsymbol{\kappa},\mathbf{p})$ or $R(\boldsymbol{\pi},\mathbf{p})$: for instance, the map $p=a\pi, q=a^{-1}\boldsymbol{\kappa}$ is homogeneous in (q,κ) variables (as $pdq=\pi d\kappa$) and it is generated by $\Psi(p,\kappa)=a^{-1}p\kappa$!
(3) A very interesting homogeneous canonical mapping is met in the theory of the motion of a rigid body (see Problems to §4.11).

Proof. If $\boldsymbol{\pi}\cdot d\boldsymbol{\kappa}-\mathbf{p}\cdot d\mathbf{q}$ is an exact differential, one sees, by going through the proof of Proposition 11, that Eq. (3.11.56) can be satisfied by choosing $H=H'$. mbe

Observations.
(1) From the proof of Proposition 22 and from Eqs. (3.11.41), (3.11.44), (3.11.47), and (3.11.49), we see that a sufficient condition in order that any Hamiltonian H on V is conjugated to a Hamiltonian H' on W given by

$$H'(\boldsymbol{\pi},\boldsymbol{\kappa},t) = H(\mathcal{C}^{-1}(\boldsymbol{\pi},\boldsymbol{\kappa},t)) \tag{3.11.66}$$

is that the transformation $\mathcal{C}$ mapping V onto W be generated by a time-independent function F or Φ or Ψ or R or be homogeneous in the sense of Proposition 22.
(2) The interest in canonical transformations consists of the fact that sometimes it is possible to solve the Hamiltonian equations by finding a canonical transformation transforming the system of Hamiltonian equations into a conjugate system with "trivial" Hamiltonian H', i.e., trivially soluble (e.g., $H'=0$ or $H'(\boldsymbol{\pi},\boldsymbol{\kappa})=h(\boldsymbol{\kappa})$ which yield trivial Hamiltonian equations, indeed).

A concrete method to look for such a transformation ("Hamilton-Jacobi method") consists in trying to find, using Proposition 21, a function F defined in a suitable neighborhood $\Omega\subset\mathcal{R}^{2\ell+1}$ such that, $\forall(\mathbf{q},\boldsymbol{\kappa},t)$,

$$H' = H(\frac{\partial F}{\partial\mathbf{q}}(\mathbf{q},\boldsymbol{\kappa},t),\mathbf{q},t) + \frac{\partial F}{\partial t}(\mathbf{q},\boldsymbol{\kappa},t) = 0 \tag{3.11.67}$$

or, for some h,

$$H' = H(\frac{\partial F}{\partial \mathbf{q}}(\mathbf{q},\boldsymbol{\kappa},t),\mathbf{q},t) + \frac{\partial F}{\partial t}(\mathbf{q},\boldsymbol{\kappa},t) = h(\boldsymbol{\kappa}) \tag{3.11.68}$$

Equations (3.11.67) and (3.11.68) are to be considered as equations in which $\boldsymbol{\kappa}$ is a parameter and, therefore, as partial differential equations for a function $(\mathbf{q},t) \to f(\mathbf{q},t)$:

$$H(\frac{\partial f}{\partial \mathbf{q}}(\mathbf{q},t),,\mathbf{q},t) + \frac{\partial f}{\partial t}(\mathbf{q},t) = 0 \tag{3.11.69}$$

or

$$H(\frac{\partial f}{\partial \mathbf{q}}(\mathbf{q},t),\mathbf{q},t) + \frac{\partial f}{\partial t}(\mathbf{q},t) = \text{constant} \tag{3.11.70}$$

("Hamilton-Jacobi" equations). We wish to find solutions to Eq. (3.11.69) or Eq. (3.11.70) which depend on ℓ parameters $\boldsymbol{\kappa} = (\kappa^{(1)},\ldots,\kappa^{(\ell)})$.

If we were able to find such a family, i.e., if we were able to find a C^∞ solution F of Eq. (3.11.69) or Eq. (3.11.70) depending on $(\mathbf{q},\boldsymbol{\kappa},t) \in \Omega = \{$ some open set in $\mathcal{R}^{2\ell+1}\}$, we could consider the transformation (3.11.41) and hope that it defines a canonical map $\mathcal{C}_F$ of some open set $\widetilde{V} \subset V$ into a set W: the transformation $\mathcal{C}_F$ would then transform the Hamiltonian equations associated with H into trivial Hamiltonian equations in W, with Hamiltonian function 0 or $h(\boldsymbol{\kappa})$.

However, it is obvious that the difficulty of solving Eqs. (3.11.67) and (3.11.68) in the above sense is equivalent to or harder than solving the original Hamiltonian equations, and one should not think of Eq. (3.11.67) or Eq. (3.11.68) as a miraculous equation.

The usefulness of the above discussion on Hamilton-Jacobi equations consists of the possibility of finding approximation algorithms to the solutions to Eq. (3.11.67) or Eq. (3.11.68) and, therefore, to the original Hamiltonian equations, which are essentially different from the general recursive method seen in §2.3, valid for solving the most general first-order differential equations.

The methods devised to construct recursively successive approximations to Eq. (3.11.67) or Eq. (3.11.68) are methods in which the particular structure of the Hamiltonian equations is explicitly used. It is therefore not too surprising that they reveal themselves to be quite appropriate to the analysis of such equations and provide better approximations for a given amount of formal work done.

The reader can convince himself of the truth of the above statement only by seeing some concrete problems studied on the basis of approximation algorithms to the solutions of the Hamilton-Jacobi equations. The best known and most celebrated of these methods or some of its variants can be found in the theory of the motion of heavenly bodies and, more generally, in the stability theory of the motion of conservative systems. An important example will be illustrated in §5.9–§5.12. Some "trivial" examples can be found in the upcoming problems.

3.11.1 Exercises, Problems and Complements

1. Construct the canonical transformation with generating function $f(q,\kappa) = \frac{m}{2}\omega q^2 \tan\kappa$, $q \in \mathcal{R}, \kappa \in \mathcal{T}^1$, and note that the above transformation simplifies the Hamiltonian $H(p,q) = \frac{p^2}{2m} + \frac{\omega^2 m}{2} q^2$. Find the harmonic oscillator motion with the help of this transformation.

2. Consider a one-dimensional mechanical system consisting of a point with mass m subject to a force with potential energy $V \in C^\infty(\mathcal{R})$. Assume that $V(0) = 0$, $V'(q) \neq 0$ if $q \neq 0$, $V(q) \xrightarrow[|q|\to+\infty]{} +\infty$. Consider the canonical map $(p,q) \to (E,\tau)$ with generating function, see p. 222, $f(E,q) = \int_0^q \sqrt{2m(E - V(q'))}\, dq'$ near a point $(\overline{p},\overline{q})$ where $\frac{\overline{p}^2}{2m} + V(\overline{q}) > 0$. Write it explicitly, finding the Hamiltonian in the new coordinates and the physical interpretation of the τ and E coordinates. (*Hint:* Do not try to "compute" the integral, but rather perform the necessary differentiations on the integral and then use the formulae for the one-dimensional motions found in §2.7).

3. Interpret f defined in Problem 2 as a solution of the equation $\frac{m}{2}\left(\frac{\partial f}{\partial q}\right)^2 + V(q) = E$ and interpret this as a one-parameter solution of the Hamilton-Jacobi equation for the mechanical system in Problem 2, in the sense of Eq. (3.11.67), of the form $f(E,q) - Et$ (or, in the sense of Eq. (3.11.68), of the form $f(E,q)$ with $h(\kappa) = E$).

4. In the context of Problem 2, define, for $E > 0$,

$$\omega(E) = \frac{\pi}{\int_{q_-(E)}^{q_+(E)} dq' \sqrt{\frac{2}{m}(E - V(q'))}},$$

where $q_\pm(E)$ are the roots of $E - V(q) = 0$. For $E > 0$, let $a(E) = \int_{E_0}^{E} \frac{dE'}{\omega(E')}\, dE'$ and let $A \to e(A)$ be its inverse function (such that $e(a(E)) \equiv E$). Consider the canonical transformation $(p,q) \to (A,\varphi)$ with generating function $\mathcal{S}(A,q) = \int_{q_0}^{q} \sqrt{2m(e(A) - V(q'))}\, dq'$ near some $(p,q) \neq (0,0)$.
Compute the new Hamiltonian and show that the canonical map may be extended to a canonical map of $\mathcal{R}^2/(0,0)$ into $(0, \int_{E_0}^{+\infty} \frac{dE'}{\omega(E')}) \times \mathcal{T}^1$. (*Hint:* Let $\varphi = \frac{\partial \mathcal{S}}{\partial A}(A,q) \mod 2\pi$, $p = \frac{\partial \mathcal{S}}{\partial a}(A,q)$ and show that this is a C^∞ map between the indicated sets.)

5. Show that the transformation in Problem 4 is a natural generalization of the Cartesian-polar coordinates in the plane (*Hint:* Consider the special case $(p^2 + q^2)/2$, where it gives exactly the Cartesian-polar coordinates. Draw the curves $A = const$ and compare them with the circles.)
The angle defined in Problem 4 is called the "average anomaly" and, therefore, the time evolution of the average anomaly is always a uniform rotation.

6. Let A, B, C be $\ell \times \ell$ matrices and A, C be symmetric. Define on $\mathcal{R}^{2\ell}$

$$F(\mathbf{q}, \boldsymbol{\kappa}, t) \stackrel{def}{=} \frac{1}{2} A\mathbf{q}\cdot\mathbf{q} + \frac{1}{2} C\boldsymbol{\kappa}\cdot\boldsymbol{\kappa} + B\boldsymbol{\kappa}\cdot\mathbf{q}.$$

Show that if $\det B \neq 0$ the map $\mathcal{C}_F$ is well defined and completely canonical between $\mathcal{R}^{2\ell+1}$ and itself. Show that its Jacobian determinant is 1, at least in the case $\ell = 1$ (the case $\ell > 1$ is discussed in §3.12). (*Hint:* First deal in detail with the case $\ell = 1$ when A, B, C are simply numbers.)

7. In the context of Problem 6, show that $\det B \neq 0$ is a necessary and sufficient condition for $\mathcal{C}_F$ to be defined. Hence, $F(q,\kappa,t) = (\kappa^2 + q^2)/2$ does not define a canonical transformation.

8. Let F be as in Problem 6. Construct explicitly the other generating functions for the canonical map [Eqs. (3.11.58), (3.11.59), and (3.11.60)] and check that, via Eqs. (3.11.44),

(3.11.46), and (3.11.49), they all generate the same completely canonical transformation if $\det A, \det B, \det C \neq 0$. Check that all the inversions mentioned in connection with the quoted formulae can actually be performed, in the present situation.

9.* Let F be as in Proposition 21. Let $(\mathbf{p}_0, \mathbf{q}_0, t_0), (\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0, t_0)$ be two points related by Eq. (3.11.41). Define the $\ell \times \ell$ matrices

$$A_{ij} = \frac{\partial^2 F}{\partial q_i \partial q_j}(\mathbf{q}_0, \boldsymbol{\kappa}_0, t_0), \quad B_{ij} = \frac{\partial^2 F}{\partial \kappa_i \partial q_j}(\mathbf{q}_0, \boldsymbol{\kappa}_0, t_0), \quad C_{ij} = \frac{\partial^2 F}{\partial \kappa_i \partial \kappa_j}(\mathbf{q}_0, \boldsymbol{\kappa}_0, t_0).$$

Show that if $\det B \neq 0$ then the map $\mathcal{C}_F$ is defined in a neighborhood of $(\mathbf{p}_0, \mathbf{q}_0, t_0)$ (*Hint:* Use Problem 6 and apply the implicit function theorem to take into account that F no longer has constant second derivatives as in the cases of Problems 6 and 8.)

10.* Is it possible that $\mathcal{C}_F$ exists near $(\mathbf{p}_0, \mathbf{q}_0, t_0)$ in the context of Problem 9 when $\det B = 0$? (*Answer:* No; hence, $F(q, \kappa, t) = f(q) + g(\kappa)$ does not define a canonical transformation. Check this directly.)

11. Show that the invertibility properties of the matrices A, B, C mentioned in connection with the quotation of Eqs. (3.11.58), (3.11.59), and (3.11.60) in Problem 8 are necessary, in general, in order to be able to express C_F as C_Φ, C_Ψ or C_R. (*Hint:* Consider, for $\ell = 1$, $F(q, \kappa) = q\kappa$: this is a case where $A = C = 0$ and the inversion cannot be realized. In this case, it is impossible to generate the corresponding canonical transformation with a function $\Psi(p, \kappa, t)$, since the transformation is easily checked to be homogeneous with respect to (κ, p) as $q dp + \pi d\kappa = 0$. See Proposition 22, p.224, and the subsequent Observation (1). Similar considerations hold for $\kappa^2 + \kappa q$, as $q dp + \pi d\kappa = -\kappa d\kappa$, which is equally impossible for reasons similar to those used in observation (1) to Proposition 22.)

12. Consider $\mathbf{x} \in \mathcal{M}_{t_1, t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ and $\mathbf{y} \in \mathcal{V}_\mathbf{x}$. Call the variation $\mathbf{y}$ "nontrivial" if $t \to \mathbf{z}(t) = \frac{\partial \mathbf{y}}{\partial \varepsilon}(t, 0)$, $t \in [t_1, t_2]$, is such that $\mathbf{z} \neq \mathbf{0}$. Define $\mathbf{x}$ to be a "strict local minimum" for the action A relative to $\mathcal{M}$ if for every variation $y \in \mathcal{V}_\mathbf{x}(\mathcal{M})$ which is nontrivial, there exists $\eta_y > 0$ such that $A(y_\varepsilon) > A(\mathbf{x})$, $\forall |\varepsilon| < \eta$, or if $\mathcal{V}_\mathbf{x}(\mathcal{M})$ only contains trivial variations. Examine the proof of Proposition 38, §2.24.1, 132, to show that in the statements of Proposition 38, §2.24.1, Proposition 6, §3.3, p.152, Proposition 8, §3.5, p.163, one can replace the words "local minimum" by "strict local minimum" (*Hint:* Just look at the proof of Proposition 38, p. 132, and Eq. (2.24.33).)

13. Let $t \to \mathbf{x}(t)$, $t \in [0, T]$, be a motion verifying the equations associated with the Lagrangian (3.11.23) and taking place in the open set $U_0 \subset \mathcal{R}^\ell$. Let E be the energy of $\mathbf{x}$ [see Eq. (3.11.33)]. Consider $\mathbf{x}$ for $t \in [t_1, t_2] \subset [0, T]$ and fix t_1: so $\mathbf{x} \in \mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E) =$ {space of the motions in $\mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2))$ taking place in U_0 and with energy E}.
Show that $\mathbf{x}$ makes stationary and (if t_2 is close enough to t_1) strictly locally minimal (see Problem 12) the action

$$\widetilde{A}(\mathbf{x}) = \int_{t_1}^{t_2} T(t)\, dt$$

in $\mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E)$. (*Hint:* Simply note that if A is stationary or strictly locally minimal on $\mathbf{x}$ in $\mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2))$ it is such in any $\mathcal{M} \subset \mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2))$. Then remark that $A(\mathbf{x}') = \int_{t_1}^{t_2} (T(t) - V(t))\, dt = \widetilde{A}(\mathbf{x}') - E\,(t_2 - t_1)$ if $\mathbf{x}' \in \mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E)$, as $T(t) + V(t) \equiv E$.)

14. Show through examples that it is possible that the set $\mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E)$ considered in Problem (13) contains finitely many points (hence $\mathcal{V}_\mathbf{x}(\mathcal{M}_{t_1, t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E))$ only contains trivial variations). Nevertheless, even in such cases the statement of Problem 12 is not an empty one: for instance, deduce from Problem 12 that the free motion in $\mathcal{R}^d$ takes places along straight lines. (*Hint:* Let $t \to \mathbf{x}(t)$ be a free motion in $\mathcal{R}^d$, then $T(t) = \frac{1}{2}\dot{\mathbf{x}}(t)^2$.

If $t \to \mathbf{x}(t)$ were not a straight line, then $\mathcal{V}_{\mathbf{x}}(\mathcal{M}_{t_1,t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E))$ would not consist only of trivial variations: however, $\widetilde{A}(\mathbf{x}') \equiv E\,(t_2 - t_1)$ for all $\mathbf{x}' \in \mathcal{M}_{t_1,t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); E)$ and, therefore, $\mathbf{x}$ could not be a strict local minimum!)

15. On $\mathcal{R}^{Nd}$, consider the metric associated with the scalar product of Eq. (3.7.1) and let $d\ell$ be the line element of the curve in $\mathcal{R}^{Nd}$ with equations $t \to \mathbf{x}(t)$, $t \in [t_i, t_2]$, which is a motion of energy E of N points, with masses $m_1, \ldots, m_N > 0$, under the influence of a force with potential energy $V \in C^\infty(\mathcal{R}^{Nd})$. Show that

$$A(\mathbf{x}) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} \sqrt{2(E - V(\boldsymbol{\xi}(\ell)))}\, d\ell - E\,(t_2 - t_1) \quad \text{if} \quad d\ell = \sqrt{2T(t)}\, dt,$$

where $A(\mathbf{x}) = \int_{t_1}^{t_2} (T(t) - V(t))\, dt$ and $\ell \to \boldsymbol{\xi}(\ell)$ is the description of the trajectory of $\mathbf{x}$ in terms of the curvilinear abscissa ℓ on it and the integral $\int_{\mathbf{x}_1}^{\mathbf{x}_2}$ is the curvilinear integral on the trajectory.

16. Consider N points in $\mathcal{R}^d$, with masses $m_1, , \ldots, m_N$. Assume that such a system is subject to an active force with potential energy $V^{(a)}$ and to an ideal holonomous constraint to a regular ℓ-dimensional surface $\Sigma \subset \mathcal{R}^{Nd}$. On Σ consider two points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ and the set, $\mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$, of the C^∞ curves joining $\boldsymbol{\xi}_1$ to $\boldsymbol{\xi}_2$ and parameterized by some parameter τ varying between 0 and 1. Given $E \in \mathcal{R}$, define on $\mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$ the curvilinear integral on the curve $\widehat{\mathbf{x}} \in \mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$ as

$$S(\widehat{\mathbf{x}}) = (\widehat{\mathbf{x}}) \int_{\boldsymbol{\xi}_1}^{\boldsymbol{\xi}_2} \sqrt{(E - V(\boldsymbol{\xi}))}\, ds,$$

where ds is the line element on E, measured with the kinetic energy metric $ds^2 = \sum_{i=1}^N m_i\, (d\mathbf{x}^{(i)})^2$, Eq. (3.7.1).
Show that the least-action principle implies that S is stationary on the curve $\widehat{\mathbf{x}}$ if and only if $\widehat{\mathbf{x}}$ is a trajectory of a motion with energy E leaving $\boldsymbol{\xi}_1$, and reaching $\boldsymbol{\xi}_2$, ("Maupertuis' principle"). (*Hint:* Consider a local system of local coordinates near Σ permitting representation of the points of $\Sigma \cap U$ through some parametric equations $\boldsymbol{\xi} = \mathbf{x}(a)$, $\mathbf{a} = (a_1, \ldots, a\ell) \in \Omega \subset \mathcal{R}^\ell$. Suppose, for simplicity, that $\widehat{\mathbf{x}}(t) \subset U \cap \Sigma$, $\forall t \in [t_1, t_2]$. Assume $\mathcal{L}$ to be a Lagrangian of the form of Eq. (3.11.23) describing the system in these coordinates. Write the stationarity conditions of S in $\mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$ for the curve $\widehat{\mathbf{x}}$ with parametric equations $\tau \to \widehat{\mathbf{a}}(\tau)$, $\tau \in [0,1]$, in the chosen coordinates. Then, in the resulting Euler-Lagrange equations, perform the change of coordinates $\tau \longleftrightarrow t$:

$$t = \int_0^\tau \sqrt{\frac{\sum_{i,j=1}^\ell g_{ij}(\mathbf{a}(\theta)) a_i'(\theta) a_j'(\theta)}{2(E - V(\mathbf{x}(\mathbf{a}(\theta))))}}\, d\theta,$$

where the prime denotes differentiation with respect to τ or θ. One finds that the motion $t \to \mathbf{x}(\mathbf{a}(\tau(t)))$ has energy E and verifies the Lagrangian equations for $\mathcal{L}$.
A more interesting *alternative* proof: consider a variation $\tau \to \widetilde{\boldsymbol{\xi}}(\tau)$ of the path $\widehat{\mathbf{x}}$ in $\mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$ and imagine it run at constant energy $\widetilde{E}$ chosen so that starting at time t_1 in $\boldsymbol{\xi}_1$ the point $\boldsymbol{\xi}_2$ is reached at time t_2 ($\widetilde{E}$ will differ by some δE from E): call $\widetilde{\mathbf{x}}(t)$ this motion which is a variation of $\widehat{\mathbf{x}}$ in $\mathcal{M}_{t_1,t_2}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$ and remark that its time law is determined by $t - t_1 = \int_{\boldsymbol{\xi}_1}^{\widetilde{\boldsymbol{\xi}}(t)} \frac{ds}{\sqrt{2(E - V(\widetilde{\mathbf{x}}))}}$. Then by Problem (13) the action of $\widetilde{\mathbf{x}}$ is

$$A(\widetilde{\mathbf{x}}) = 2 \int_{t_1}^{t_2} \widetilde{T}(t) dt - \widetilde{E}(t_2 - t_1) = \int_{t_1}^{t_2} \sqrt{2\widetilde{T}(t)} \sqrt{2\widetilde{T}(t)} dt - \widetilde{E}(t_2 - t_1)$$

and the latter expression can be written $(\widetilde{\mathbf{x}}) \int_{\boldsymbol{\xi}_1}^{\boldsymbol{\xi}_2} \sqrt{2} \sqrt{\widetilde{E} - V(\widetilde{\mathbf{x}}(s))} ds - \widetilde{E}(t_2 - t_1)$. Thus

$$\delta A = \sqrt{2}\delta S + \int_{\boldsymbol{\xi}_1}^{\boldsymbol{\xi}_2} \frac{ds}{\sqrt{2(E - V(\widetilde{\mathbf{x}}(s)))}} \delta E - (t_2 - t_1)\,\delta E \equiv \sqrt{2}\delta S$$

hence stationarity of S on a curve in $\mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2|\Sigma)$ implies stationarity of A on the corresponding motion running on the curve with energy E.)

17. In the context of Problem 16, show that the Maupertuis' principle can be interpreted as saying that the motions developing on Σ with energy E take place along the geodesics of Σ with respect to the metric on Σ:

$$dh = \sqrt{2(E - V(\boldsymbol{\xi}))}\, ds,$$

where ds the kinetic energy metric on Σ. (We recall that by definition, a curve on Σ is called a geodesic for a given line element on Σ if it makes stationary the distance between any two of its points measured along the curve itself using the given line element.)
In other words, if we call the distance between two points $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Sigma$ measured with the line element dh along a given curve on Σ, with the name "mechanical path" with energy E on Σ, we can say that the "motions with energy E on Σ take place along trajectories making stationary the mechanical path with energy E". As usual, it is possible to show that the mechanical systems with Lagrangian Eq. (3.11.23) have the property that a trajectory of any of their motions with energy E, taking place on Σ, not only makes the mechanical path stationary but actually strictly minimizes it on short enough segments.

18.* Under the assumptions of Problem 17, let $s \to \widehat{\mathbf{x}}(s), s \in [s_1, s_2]$, be a geodesic segment on Σ for the line element dh. Suppose that $E - V(\widehat{\mathbf{x}}(s)) > 0, \forall s \in [s_1, \widehat{s}]$. Show that there is $\overline{s} > s_1$, such that if $s_2 \in [s_1, \overline{s}]$, the curve $s \to \widehat{\mathbf{x}}(s)$, $s \in [s_1, s_2]$, makes strictly locally minimal the mechanical path with energy E between $\widehat{\mathbf{x}}(s_1)$ and $\widehat{\mathbf{x}}(s_2)$.

19. A point with mass $m > 0$ is bound to a surface $\Sigma \subset \mathcal{R}^3$ by an ideal constraint and it is subject to no other forces. Show that as a consequence of the Maupertuis principle, Problems 16-18, the point runs on Σ in such a way that if two points on its trajectory are close enough, then the trajectory itself is the one minimizing the distance on Σ between the two points, i.e., the trajectory is the shortest path on Σ joining the two points, the distance being measured in the ordinary $\mathcal{R}^3$ sense ("geodesics" or Fermat's principle"). (*Hint:* Note that dh and ds are now proportional, and use Problem 18.)

20. Consider the line segment $(dx^2 + dy^2)/y^2$ defined on the half-plane $y > 0$. Determine its geodesics by thinking of them via the mechanical interpretation, permitted by Problem 16, which allows us to regard them as the zero energy motions of the mechanical system with Lagrangian $\mathcal{L} = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{2y^2}$.

21. Calling the geodesics of the Problem 20 "straight lines for the geometry ("Lobachevski geometry" or "noneucidean geometry") defined by the line element ds", check the truth or the falsity of the following statements:
(i) Given two points in the half-plane $y > 0$, there is one and only one straight fine through them.
(ii) Two points in the $y > 0$ region are joined by just one straight line segment (if a straight line segment is defined as a connected closed subset of a straight line).
(iii) Given a point, and a straight line not containing it, there exists just one straight line containing the point and "parallel" to the first straight line (i.e., without common points with it).

22. Same as Problems 20 and 21 for the geometries associated with the following line elements:

$(i) \qquad ds^2 = (x^2 + y^2)(dx^2 + dy^2), \qquad (x, y) \in \mathcal{R}^2 \backslash \mathbf{0};$

$(ii) \qquad ds^2 = (1 - x^2 - y^2)^\alpha (dx^2 + dy^2), \qquad (x, y) \in \mathcal{R}^2,\ x^2 + y^2 < 1, \alpha \in \mathcal{R};$

$(iii) \qquad ds^2 = \dfrac{dx^2 + dy^2}{\sqrt{x^2 + y^2}}, \qquad (x, y) \in \mathcal{R}^2 \backslash \mathbf{0}.$

23. Same as Problems 20 and 21 for the geometry defined on a sphere by the line element induced by the Euclidean distance of $\mathcal{R}^\ell$; i.e., $ds^2 = d\theta^2 + \sin^2\theta d\varphi^2$ in polar coordinates.

24. Consider the geometry defined in the half-plane $y > 0$ by the line element of Problem 20. Define a "triangle" as a figure formed by the three points pairwise connected by geodesic segments. Given a triangle, denote α, β, γ the three angles relative to its three vertices (defined as the angles between the tangents to the two geodesic segments meeting at the various vertices). The quantity $\alpha + \beta + \gamma - \pi$ is called the "geodesic defect": show that it is < 0. Show that the same quantity computed in the analogous situation for the sphere's geometry of Problem 23 is > 0.

25. A light ray moves in a plane strip $x \in \mathcal{R}, |y| < 1$ with refraction index

$$n(x, y) = \sqrt{1 - \varepsilon y^2}, \qquad \varepsilon < 1.$$

Using Fermat's principle, show that the ray proceeds along a sinusoidal path, if it is assumed that the ray starts at the origin with an initial direction close to the horizontal. Recall, for this purpose, that Fermat's principle says that the rays follow a path that makes stationary the "optical path" between any two of its points, within the set of the paths joining them. The optical path, in a medium with index of refraction $n(x, y)$, associated with the curve $\widehat{\mathbf{x}} \in \mathcal{M}_{0,1}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$, is

$$(\widehat{\mathbf{x}}) \int_{\boldsymbol{\xi}_1}^{\boldsymbol{\xi}_2} n(x, y) ds, \quad ds = \sqrt{dx^2 + dy^2}.$$

(*Hint:* Interpret the above problem as a mechanical problem via Problems 16 and 17.)

Hence, via Maupertuis' principle, the problem of the determination of a light path can be interpreted as a purely mechanical problem.

26. Solve the problems at the end of §18, §20, §21, §24, §32,§39,§44 in [28].

27. Perform the the Legendre transformation on the Lagrangian $\mathcal{L} = \sqrt{\dot{x}^2 + \dot{y}^2}$ and explain why one gets strange results.

28. Consider the function on $\mathcal{R}^{2\ell}$: $\mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}) = \frac{1}{2} A\dot{\mathbf{q}} \cdot \dot{\mathbf{q}} + \frac{1}{2} C\mathbf{q} \cdot \mathbf{q} + B\dot{\mathbf{q}} \cdot \mathbf{q}$ where A, C are $\ell \times \ell$ symmetric matrices and B is an $\ell \times \ell$ matrix. Under which conditions on A, B, C is $\mathcal{L}$ a regular Lagrangian on $\mathcal{R}^{2\ell}$? In these cases, write the corresponding Hamiltonian function. Similarly, consider the function on $\mathcal{R}^{2\ell}$: $H(\mathbf{p}, \mathbf{q}) = \frac{1}{2} A\mathbf{p} \cdot \mathbf{p} + \frac{1}{2} C\mathbf{q} \cdot \mathbf{q} + B\mathbf{p} \cdot \mathbf{q}$ and find the conditions for H to be a regular Hamiltonian and write the corresponding Lagrangian.

29. In the cases when the Lagrangian in Problem 28 is regular, write the energy conservation theorem, Proposition 18, §3.11, in terms of $\dot{\mathbf{q}}$ and $\mathbf{q}$. (*Hint:* $H = \mathbf{p} \cdot \mathbf{q} - \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q})$, and then express $\mathbf{p}$ in terms of $\dot{\mathbf{q}}, \mathbf{q}$ and use Proposition 18.)

30. Show that the time-independent completely canonical linear transformations of $\mathcal{R}^{2\ell}$ onto $\mathcal{R}^{2\ell}$ form a group $\mathcal{S}_\ell$, under the natural composition law.

31. The set $\mathcal{G}$ of the linear completely canonical transformations on $\mathcal{R}^2$ with generating functions $\Phi(\pi, q) = \frac{1}{2} a\pi^2 + \frac{1}{2} cq^2 + b\pi q$, $b \neq 0$, which we denote (a, c, b), does not form a subgroup of $\mathcal{S}_1$. Prove this by finding the composition law of (a, c, b) and (a', c', b').

Show that $(a,c,b)\cdot(a',c',b')\in\mathcal{G}$ if and only if $a'c\neq 1$. (*Hint:* The composition law, if $\delta=ac-b^2,\delta'=a'c'-\delta^2$, is

$$(a,c,b)\cdot(a',c',b')=\left(\frac{a-a'\delta}{1-a'c},\frac{-c'+c\delta'}{1-a'c},\frac{bb'}{1-a'c}\right).)$$

32. Same as Problem 31 for the class $\mathcal{G}'$ of the canonical transformations generated by functions $F(\kappa,q)=\frac{1}{2}aq^2+\frac{1}{2}c\kappa^2+bq\kappa$, $b\neq 0$. (*Hint:* The composition law is now, for suitable δ,δ'

$$(a,c,b)\cdot(a',c',b')=\left(\frac{aa'-\delta'}{a+c'},\frac{cc'-\delta}{a+c'},\frac{-bb'}{a+c'}\right).)$$

33. Find a generating function for the transformation $(\boldsymbol{\pi},\boldsymbol{\kappa})\to(\mathbf{p},\mathbf{q})$ defined by $\mathbf{p}=R\boldsymbol{\pi},\mathbf{q}=(R^T)^{-1}\boldsymbol{\kappa}$, where R is a nonsingular $\ell\times\ell$ matrix: this transformation is completely canonical. (*Hint:* Look for a generating function like $\Phi(\boldsymbol{\pi},\mathbf{q})=B\boldsymbol{\pi}\cdot\mathbf{q}$ with B being an $\ell\times\ell$ matrix.)

34. Let $\boldsymbol{\kappa}\to\mathbf{q}=\mathbf{f}(\boldsymbol{\kappa})$ be an invertible nonsingular transformation of $\mathcal{R}^\ell$ onto itself. Find out how to define $\mathbf{p}=\mathbf{F}(\boldsymbol{\pi},\boldsymbol{\kappa})$ so that the map $(\boldsymbol{\pi},\boldsymbol{\kappa})\to(\mathbf{p},\mathbf{q})=(F(\boldsymbol{\pi},\boldsymbol{\kappa}),\mathbf{f}(\boldsymbol{\kappa}))$ will be completely canonical. (*Answer:* If $R_{ij}(\boldsymbol{\kappa})=\frac{\partial f_i}{\partial\kappa_j}(\boldsymbol{\kappa})$, then $\mathbf{p}=(R(\boldsymbol{\kappa})^T)^{-1}\boldsymbol{\pi}$.)

35. Let $f\in C^\infty(\mathcal{R}^\ell)$ be multi periodic with periods 2π. Is the function $\mathbf{A}'\cdot\boldsymbol{\varphi}+f(\boldsymbol{\varphi})$ a generating function of a canonical map of $\mathcal{R}^\ell\times\mathcal{T}^\ell$ onto $\mathcal{R}^\ell\times\mathcal{T}^\ell$? Find a sufficient condition.

36. Let $(\mathbf{A}',\boldsymbol{\varphi})\to f(\mathbf{A}',\boldsymbol{\varphi})$ be a $C^\infty(\mathcal{R}^{2\ell})$ function multi periodic with periods 2π in the $\boldsymbol{\varphi}$'s. Suppose that the transformation

$$\boldsymbol{\varphi}'=\boldsymbol{\varphi}+\frac{\partial f}{\partial\mathbf{A}'}(\mathbf{A}',\boldsymbol{\varphi})\mod 2\pi$$

establishes a nonsingular invertible map of $\mathcal{T}^\ell$ onto itself for each $\mathbf{A}'\in\mathcal{R}^\ell$. Suppose, also, that the transformation

$$\mathbf{A}=\mathbf{A}'+\frac{\partial f}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi})$$

establishes a nonsingular invertible map of $\mathcal{R}^\ell$ onto itself for each $\boldsymbol{\varphi}\in\mathcal{T}^\ell$.
Show that the function $\Phi(\mathbf{A}',\boldsymbol{\varphi})=\mathbf{A}'\cdot\boldsymbol{\varphi}+f(\mathbf{A}',\boldsymbol{\varphi})$ generates a completely canonical map of $\mathcal{R}^\ell\times\mathcal{T}^\ell$ onto itself.

37. Find a "local version" of Problem 36 when $\mathcal{R}^\ell\times\mathcal{T}^\ell$ is replaced by $V\times\mathcal{T}^\ell$, $V\subset\mathcal{R}^\ell$ open.

38. Consider the maps $(p,q)\to\mathcal{C}(p,q)=(A,\varphi)$ and $(p,q)\to\widetilde{\mathcal{C}}(p,q)=(B,\varphi)$ of $\mathcal{R}^2/\mathbf{0}\longleftrightarrow\mathcal{R}_+\times\mathcal{T}^\ell$ defined by

$$\varphi=\text{polar angle of }(p,q),\qquad A=\frac{1}{2}(p^2+q^2),\ B=\sqrt{A}.$$

Show that while $\mathcal{C}$ is completely canonical, the map $\widetilde{\mathcal{C}}$ is such that the Hamiltonians $H=\frac{1}{2}(p^2+q^2)$ and $H'=B$ are canonically conjugated by it, but the Hamiltonian $\frac{1}{2}p^2$ has no canonically conjugated Hamiltonian with respect to $\widetilde{\mathcal{C}}$. (*Hint:* $\mathcal{C}$ is studied in Problems 1 and 2. Show that a general measure-preserving flow on $\mathcal{R}^2/\mathbf{0}$ is not mapped by $\widetilde{\mathcal{C}}$ into a measure-preserving flow on $\mathcal{R}_+\times\mathcal{T}^\ell$: the evolution associated with $H'=\frac{1}{2}p^2$ is actually mapped by $\widetilde{\mathcal{C}}$ into a non-measure preserving one. So the image flow cannot be a Hamiltonian flow since the latter would, instead, preserve the measure, by the Liouville theorem, Proposition

19, in the case of a time-independent Hamiltonian (or by an extension of Proposition 19 in the time-dependent case; see Problem 39).)

39. Extend Proposition 19, §3.11 to the case of time-dependent Hamiltonian equations. (*Hint:* Replace the semigroup property $S_t S_{t'} = S_{t+t'}$ used in the proof of proposition 19 (see Problem 10, §2.24) by the more general relation $S(t,t')\cdot S(t',t_0) = S(t,t_0)$, $t > t' > t_0$, where $S(t,t')$ denotes the solution map of the non autonomous Hamiltonian equations when the initial data are assigned at t'. The proof proceeds unchanged.)

40. Let $(\mathbf{q}, t) \to S(\mathbf{q}, t)$ be defined and C^∞ on a set $U \times J$, $U \subset \mathcal{R}^\ell$, $J \subset \mathcal{R}$, both open, and connected. Let H be a regular Hamiltonian on $V = \mathcal{R}^\ell \times U \times J$ and suppose that S is a solution to the Hamilton-Jacobi equation

$$H(\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, t), \mathbf{q}, t) + \frac{\partial S}{\partial t}(\mathbf{q}, t) = 0.$$

Consider the differential equation for $t \to \mathbf{q}(t)$,

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}\Big(\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, t), \mathbf{q}, t\Big), \quad \mathbf{q}(t_0) = \mathbf{q}_0$$

and suppose that for all $(\mathbf{q}_0, t_0) \in U \times J$, one can solve it for t near t_0 by $t \to \mathbf{q}(t)$. Show that setting

$$\mathbf{p}(t) = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}(t), t),$$

the functions $t \to (\mathbf{p}(t), \mathbf{q}(t))$ are solutions to the Hamiltonian equations with initial data

$$\mathbf{q}(t_0) = \mathbf{q}_0, \quad \mathbf{p}(t_0) = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}_0, t_0);$$

i.e., "every solution to the Hamilton-Jacobi equation provides a bundle of solutions to the corresponding Hamilton equation". (*Hint:* Check it directly by substitution.)

3.12 Completely Canonical Transformations: Their Structure

Among the canonical transformations, the completely canonical transformations are very simple and interesting [see Eq. (3.11.39)]. It is therefore important to obtain general results about the structure of such maps.

Let $V \subset \mathcal{R}^\ell \times \mathcal{R}^\ell$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2})$, $\ell_1 + \ell_2 = \ell$, be an open set which is regarded as the phase space of a Hamiltonian systems of differential equations with regular Hamiltonian functions $H \in C^\infty(V)$ (see Observation (4) to Proposition 17, p.216.)

18 Definition. *Let V, W be open sets as above and let $\mathcal{C}$ be an invertible nonsingular*[21] *C^∞ map between V and W. Denote $\mathcal{C}$*

$$\mathbf{p} = \mathbf{P}(\boldsymbol{\pi}, \boldsymbol{\kappa}), \qquad \mathbf{q} = \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa}) \tag{3.12.1}$$

[21] i.e., with non vanishing Jacobian determinant.

Let $(\mathbf{p}_0, \mathbf{q}_0) = \mathcal{C}(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0)$, $(\mathbf{p}_0, \mathbf{q}_0) \in V$, $(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0) \in W$ *be two* $\mathcal{C}$*-corresponding points. Define the "linearized* $\mathcal{C}$ *map near* $(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0)$*" as the map of* $\mathcal{R}^\ell \times \mathcal{R}^\ell$

$$\begin{aligned} \mathbf{p} &\stackrel{def}{=} \mathbf{p}_0 + A(\boldsymbol{\pi} - \boldsymbol{\pi}_0) + B(\boldsymbol{\kappa} - \boldsymbol{\kappa}_0), \\ \mathbf{q} &\stackrel{def}{=} \mathbf{q}_0 + C(\boldsymbol{\pi} - \boldsymbol{\pi}_0) + D(\boldsymbol{\kappa} - \boldsymbol{\kappa}_0), \end{aligned} , \tag{3.12.2}$$

where A, B, C *and* D *are the four* $\ell \times \ell$ *matrices:*

$$\begin{aligned} A_{ij} &\stackrel{def}{=} \frac{\partial P_i}{\partial \pi_j}(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0), \qquad & B_{ij} &\stackrel{def}{=} \frac{\partial P_i}{\partial \kappa_j}(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0), \\ C_{ij} &\stackrel{def}{=} \frac{\partial Q_i}{\partial \pi_j}(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0), \qquad & D_{ij} &\stackrel{def}{=} \frac{\partial Q_i}{\partial \kappa_j}(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0), \end{aligned} \tag{3.12.3}$$

$i, j = 1 \ldots, \ell$, *and in Eq. (3.12.2),* $\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0, \mathbf{p}_0, \mathbf{q}_0$ *are regarded as elements of* $\mathcal{R}^\ell$ *(even though* $\boldsymbol{\kappa}_0, \mathbf{q}_0$ *might be in* $\mathcal{T}^\ell$ *or* $\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}$*).*[22] *The* $2\ell \times 2\ell$ *matrix*

$$L = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{3.12.4}$$

is the Jacobian matrix of the map $\mathcal{C}$ *and, therefore,* $\det L \neq 0$, $\forall (\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0) \in W$.

The main structure theorem for the completely canonical time independent maps transforming V onto W is:

23 Proposition. *A necessary and sufficient condition for the complete canonicity of a map* $\mathcal{C}$ *of the type considered in the Definition 18, Eq. (3.12.1), is that the map obtained by linearizing* $\mathcal{C}$ *at* $(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0) \in W$ *is a completely canonical map of* $\mathcal{R}^{2\ell}$ *onto* $\mathcal{R}^{2\ell}$, $\forall (\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0) \in W$.
This is the case if and only if the inverse matrix to the matrix (3.12.4) is

$$L^{-1} = \begin{pmatrix} D^T & -B^T \\ -C^T & A^T. \end{pmatrix} \tag{3.12.5}$$

where the superscript T *denotes the transposition of the matrix.*

Observations.
(1) In other words, $\mathcal{C}$ is completely canonical in W if and only if its linearization around any point in W is completely canonical.
(2) Hence, complete canonicity is a "purely local" property of a map: this explains why the completely canonical maps are sometimes called "contact transformations" (although it does not explain why they are often called "symplectic").

PROOF. Let $H \in C^\infty(V)$ and $H'(\boldsymbol{\pi}, \boldsymbol{\kappa}) = H(\mathcal{C}(\boldsymbol{\pi}, \boldsymbol{\kappa})) = H(\mathbf{P}(\boldsymbol{\pi}, \boldsymbol{\kappa}), \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa}))$. The Hamiltonian equations in V are

[22] We recall that on $\mathcal{T}^\ell$ we use the flat coordinates: the ambiguity $\bmod 2\pi$ of some of the coordinates of $\boldsymbol{\kappa}_0$ or $\mathbf{q}_0$ is arbitrarily solved here and it is irrelevant in the following.

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}(\mathbf{p}, \mathbf{q}), \qquad \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}(\mathbf{p}, \mathbf{q}) \tag{3.12.6}$$

and if $\mathcal{C}$ is completely canonical, they must be equivalent to the equations

$$\dot{\boldsymbol{\pi}} = -\frac{\partial H'}{\partial \boldsymbol{\kappa}}(\boldsymbol{\pi}, \boldsymbol{\kappa}), \qquad \dot{\boldsymbol{\kappa}} = \frac{\partial H'}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi}, \boldsymbol{\kappa}) \tag{3.12.7}$$

i.e., if $t \to (\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t))$ solves Eq. (3.12.7), then the motion $t \to \mathcal{C}(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t)) = (\mathbf{P}(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t)), \mathbf{Q}(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t))) \equiv (\mathbf{p}(t), \mathbf{q}(t))$ has to solve Eq. (3.12.6). Differentiating $p_i(t) = P_i(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t)), Q_i(t) = q_i(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t))$ with respect to t,

$$\begin{aligned} \dot{p}_i &= \sum_{i=1}^{\ell} \big(A_{ik}\dot{\pi}_k + B_{ik}\dot{\kappa}_k\big) = \sum_{k=1}^{\ell} \Big(- A_{ik}\frac{\partial H'}{\partial \kappa_k} + B_{ik}\frac{\partial H'}{\partial \pi_k}\Big), \\ \dot{q}_i &= \sum_{i=1}^{\ell} \big(C_{ik}\dot{\pi}_k + D_{ik}\dot{\kappa}_k\big) = \sum_{k=1}^{\ell} \Big(- C_{ik}\frac{\partial H'}{\partial \kappa_k} + D_{ik}\frac{\partial H'}{\partial \pi_k}\Big), \end{aligned} \tag{3.12.8}$$

for $i = 1, \ldots, \ell$, where the matrices A, B, C, D and the derivatives of H' are evaluated at $(\boldsymbol{\pi}(t), \boldsymbol{\kappa}(t))$, to simplify the notations. Using the expression of H' in terms of H, we find from Eq. (3.12.8), $\forall i = 1, \ldots \ell$:

$$\begin{aligned} \dot{p}_i &= \sum_{k,s} \Big\{\Big(- A_{ik}(\frac{\partial H}{\partial p_s}B_{sk} + \frac{\partial H}{\partial q_s}D_{sk}) + B_{ik}(\frac{\partial H}{\partial p_s}A_{sk} + \frac{\partial H}{\partial q_s}C_{sk})\Big)\Big\} \\ \dot{q}_i &= \sum_{k,s} \Big\{\Big(- C_{ik}(\frac{\partial H}{\partial p_s}B_{sk} + \frac{\partial H}{\partial q_s}D_{sk}) + D_{ik}(\frac{\partial H}{\partial p_s}A_{sk} + \frac{\partial H}{\partial q_s}C_{sk})\Big)\Big\}, \end{aligned} \tag{3.12.9}$$

where the derivatives of H are computed in the point $(\mathbf{P}(\boldsymbol{\pi}, \boldsymbol{\kappa}), \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa}))$, and the matrices A, B, C, and D have to be computed in $(\boldsymbol{\pi}, \boldsymbol{\kappa})$. Equation (3.12.9) can be more compactly written with matrix-product notations:

$$\begin{pmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{q}} \end{pmatrix} = \begin{pmatrix} (AB^T - BA^T) & (-AD^T + BC^T) \\ (-DA^T + CB^T) & (-CD^T + DC^T) \end{pmatrix} \begin{pmatrix} -\frac{\partial H}{\partial \mathbf{p}} \\ \frac{\partial H}{\partial \mathbf{q}} \end{pmatrix} \tag{3.12.10}$$

We now impose that Eq. (3.12.10) reduces to Eq. (3.12.6), $\forall H \in C^\infty(V)$. Since the vector in the right-hand side of Eq. (3.12.10) can be made arbitrary by varying H, if the point $(\mathbf{p}, \mathbf{q})$ where the derivatives are evaluated is kept fixed, it follows that

$$AB^T - BA^T = 0, \quad CD^T - DC^T = 0, \quad -AD^T + BC^T = -I, \tag{3.12.11}$$

where $I = (\ell \times \ell$ identity matrix). Note that

$$\begin{pmatrix} (AB^T - BA^T) & (-AD^T + BC^T) \\ (-DA^T + CB^T) & (-CD^T + DC^T) \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} B^T & -D^T \\ -A^T & C^T \end{pmatrix} \tag{3.12.12}$$

hence, Eq. (3.12.11) can be written as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} B^T & -D^T \\ -A^T & C^T \end{pmatrix} = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} \tag{3.12.13}$$

or, multiplying both sides on the right by $\begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix}$:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} D^T & -B^T \\ -C^T & A^T \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{3.12.14}$$

which implies Eq. (3.12.5).

Vice versa, if Eq. (3.12.5) holds everywhere in W, the above equalities can be run backwards.

If H is explicitly time dependent, its conjugacy via $\mathcal{C}$ with H' defined by $H'(\boldsymbol{\pi}, \boldsymbol{\kappa}, t) = H(\mathcal{C}(\boldsymbol{\pi}, \boldsymbol{\kappa}), t)$ follows in an identical fashion. mbe

24 Proposition. *The Jacobian determinant of any completely canonical transformation is* ± 1.

PROOF. Equation (3.12.14) can be written

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} = -\mathbf{1} \tag{3.12.15}$$

where $\mathbf{1}$ denotes the $2\ell \times 2\ell$ identity matrix; i.e.,

$$L \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} L^T \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} = -\mathbf{1}. \tag{3.12.16}$$

Hence, taking the determinant of both sides and remarking that the matrix $E = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ has determinant $\det E = +1$, it follows that

$$(\det L)^2 = 1 \tag{3.12.17}$$

mbe

It could be shown that, actually, $\det L = +1$, see problem (16) below.

The conditions (3.12.15) or (3.12.11) for complete canonicity, equivalent to Eq. (3.12.5), can be expressed in terms of the following notion of "Poisson bracket" of two observables.

19 Definition. *Let V be an open subset of $\mathcal{R}^ell \times \mathcal{R}^\ell$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}), ell_l + \ell_2 = \ell$, regarded as the phase space for the Hamiltonian equations in V. Let $F, G \in C^\infty(V)$ be two "observables". One defines the "Poisson bracket" $\{F, G\} \in C^\infty(V)$ of F and G as*

$$\{F, G\}(\mathbf{p}, \mathbf{q}) = \sum_{s=1}^{\ell} \Big(\frac{\partial F}{\partial p_s}(\mathbf{p}, \mathbf{q}) \frac{\partial G}{\partial q_s}(\mathbf{p}, \mathbf{q}) - \frac{\partial F}{\partial q_s}(\mathbf{p}, \mathbf{q}) \frac{\partial G}{\partial p_s}(\mathbf{p}, \mathbf{q}) \Big). \quad (3.12.18)$$

Observations.
(1) Clearly, $\forall i, j = 1, \ldots, \ell$,

$$\{p_i, q_j\} = \delta_{ij}, \ \{p_i, p_j\} = 0, \ \{q_i, q_j\} = 0 \quad (3.12.19)$$

Also, if $\varphi_1, \ldots, \varphi_r$ are C^∞ functions on $\mathcal{R}^n$ and $F_1, \ldots, F_n \in C^\infty(V)$, and if one defines

$$\Phi_j(\mathbf{p}, \mathbf{q}) = \varphi_j(F_1(\mathbf{p}, \mathbf{q}), \ldots, F_n(\mathbf{p}, \mathbf{q})), \quad (3.12.20)$$

one finds:

$$\{\Phi_i, \Phi_j\} = \sum_{h,k=1}^{n} \frac{\partial \varphi_i}{\partial F_h} \frac{\partial \varphi_j}{\partial F_k} \{F_h, F_k\}. \quad (3.12.21)$$

(2) Sometimes the definition (3.12.18) is given with the opposite sign: this is totally irrelevant despite claims to the contrary.
(3) Equations (3.12.19) are also called the "canonical commutation" relations.

The notion of Poisson bracket is remarkable as it appears from the following corollary to the Proposition 23, p.234.

25 Corollary. *A necessary and sufficient condition for the complete canonicity of an invertible nonsingular map $\mathcal{C}$ between V and W (as in Definition 17, p.220, above) is that the functions defining it, $\mathbf{P}(\boldsymbol{\pi}, \boldsymbol{\kappa}), \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa})$, have the property, $\forall(\boldsymbol{\pi}, \boldsymbol{\kappa}) \in W$, $\forall i, j = 1, \ldots, \ell$:*

$$\{P_i, Q_j\} = \delta_{ij}, \ \{P_i, P_j\} = 0, \ \{Q_i, Q_j\} = 0 \quad (3.12.22)$$

Observations.
(1) So $\mathcal{C}$ is completely canonical if and only if it "preserves the canonical commutation relations".
(2) If $\mathcal{C}$ preserves the canonical commutation relations, it follows that it preserves the Poisson brackets of any pair of observables: this means that if $F, G \in C^\infty(V)$ and if we define

$$F_{\mathcal{C}}(\boldsymbol{\pi}, \boldsymbol{\kappa}) = F(\mathcal{C}(\boldsymbol{\pi}, \boldsymbol{\kappa})), \qquad G_{\mathcal{C}}(\boldsymbol{\pi}, \boldsymbol{\kappa}) = G(\mathcal{C}(\boldsymbol{\pi}, \boldsymbol{\kappa})), \quad (3.12.23)$$

then, as is checked by Eqs. (3.12.21) and (3.12.22):

$$\{F, G\}(\mathbf{p}, \mathbf{q}) = \{F_{\mathcal{C}}, G_{\mathcal{C}}\}(\boldsymbol{\pi}, \boldsymbol{\kappa}) \qquad \text{if } (\mathbf{p}, \mathbf{q}) = \mathcal{C}(\boldsymbol{\pi}, \boldsymbol{\kappa}). \tag{3.12.24}$$

So $\mathcal{C}$ is completely canonical if and only if it "preserves the commutation relations of any pair of observables".

PROOF. Explicitly write Eq. (3.12.22) in terms of the derivatives of Eq. (3.12.3): one finds that they become Eq. (3.12.11), i.e., Eq. (3.12.5).

mbe

In §3.11 it has been shown that a class of completely canonical transformations can be built from a generating function. One can wonder how general this construction is.

26 Proposition. *Let $\mathcal{C}$ be a completely canonical map between V and W as in Definition 17, p.220. Given two corresponding points $(\mathbf{p}_0, \mathbf{q}_0) \in V$, $(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0) \in W$, $(\mathbf{p}_0, \mathbf{q}_0) = \mathcal{C}(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0)$, consider the matrices (3.12.3). Then $\mathcal{C}$ can be generated near $(\mathbf{p}_0, \mathbf{q}_0)$, $(\boldsymbol{\pi}_0, \boldsymbol{\kappa}_0)$ by a generating function, as in Proposition 21, p.220, and in the observations following it, having the form:*

$$\begin{array}{lll} (i) & F(\mathbf{q}, \boldsymbol{\kappa}) & \text{if } \det C \neq 0, \\ (ii) & \Phi(\mathbf{p}, \boldsymbol{\kappa}) & \text{if } \det A \neq 0, \\ (iii) & \Psi(\boldsymbol{\pi}, \mathbf{q}) & \text{if } \det D \neq 0, \\ (iv) & R(\mathbf{p}, \boldsymbol{\pi}) & \text{if } \det B \neq 0, \end{array} \tag{3.12.25}$$

Observations.
(1) There exist completely canonical transformations for which $\det A = \det B = \det C = \det D = 0$. For instance, the map of $\mathcal{R}^2 \times \mathcal{R}^2 \longleftrightarrow \mathcal{R}^2 \times \mathcal{R}^2$

$$(p_1, p_2; q_1, q_2) \longleftrightarrow (p_1, -q_2; q_1, p_2). \tag{3.12.26}$$

This canonical transformation cannot be generated by a generating function of the above types.
(2) If $\mathcal{C}$ is completely canonical, defined on $V \subset \mathcal{R}^{2\ell}$, it must have a Jacobian matrix L with non vanishing determinant, (see Proposition 24, p.236). Hence, there must be a choice of indices $i_1, \ldots, i_s, j_1, \ldots j_{\ell-s}$, pairwise distinct, with

$$\det \frac{\partial(p_{i_1}, \ldots, p_{i_s}, q_{j_1}, \ldots, q_{j_{\ell-s}})}{\partial(\pi_1, \ldots, \pi_\ell)} \neq 0 \tag{3.12.27}$$

This means, as it can be understood with a little thought, that $\mathcal{C}$ can be locally constructed by composing a canonical transformation of the type:

$$\begin{aligned} &(p_1, \ldots, p_\ell; q_1, \ldots, q_\ell) \longleftrightarrow \\ &(p_{i_1}, \ldots, p_{i_s}, -q_{j_1}, \ldots, -q_{j_{\ell-s}}; q_{i_1}, \ldots, q_{i_s}, p_{j_1}, \ldots, p_{j_{\ell-s}}) \end{aligned} \tag{3.12.28}$$

[like Eq. (3.12.26)] with a completely canonical transformation generated by a function $\Phi(\mathbf{p}, \boldsymbol{\kappa})$.

(3) So any completely canonical transformation is, near a point, a composition of a trivial ("permutation type") completely canonical transformation and a completely canonical transformation with a generating function (Arnold).

PROOF. Suppose, for instance, $\det D \neq 0$. Then it is possible to invert the equation $\mathbf{q} = \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa})$ to express

$$\boldsymbol{\kappa} = \mathbf{G}(\boldsymbol{\pi}, \mathbf{q}), \tag{3.12.29}$$

for $(\mathbf{q}, \boldsymbol{\pi}, \boldsymbol{\kappa})$ near $(\mathbf{q}_0, \boldsymbol{\pi}_0, \boldsymbol{\kappa}_0)$, using the implicit function theorem (see Appendix G). Then we can write

$$\mathbf{p} = \mathbf{P}(\boldsymbol{\pi}, \mathbf{G}(\boldsymbol{\pi}, \mathbf{q})) \equiv \mathbf{F}(\boldsymbol{\pi}, \mathbf{q}), \qquad \boldsymbol{\kappa} = \mathbf{G}(\boldsymbol{\pi}, \mathbf{q}) \tag{3.12.30}$$

and we must show existence of $\boldsymbol{\Psi}(\boldsymbol{\pi}, \mathbf{q})$ such that

$$\frac{\partial \psi}{\partial \mathbf{q}}(\boldsymbol{\pi}, \mathbf{q}) = \mathbf{F}(\boldsymbol{\pi}, \mathbf{q}), \qquad \frac{\partial \psi}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi}, \mathbf{q}) = \mathbf{G}(\boldsymbol{\pi}, \mathbf{q}) \tag{3.12.31}$$

defined near $\boldsymbol{\pi}_0, \mathbf{q}_0$. This means checking the integrability conditions:

$$\frac{\partial F_i}{\partial q_j} = \frac{\partial F_j}{\partial q_i}, \quad \frac{\partial F_i}{\partial \pi_j} = \frac{\partial G_j}{\partial q_i}, \quad \frac{\partial G_i}{\partial \pi_j} = \frac{\partial G_j}{\partial \pi_i}, \tag{3.12.32}$$

Differentiation of the first of Eqs. (3.12.30) yields

$$\frac{\partial F_i}{\partial \pi_j} = A_{ij} + \sum_{k=1}^{\ell} B_{ik} \frac{\partial G_k}{\partial \pi_j}, \qquad \frac{\partial F_i}{\partial q_j} = \sum_{k=1}^{\ell} B_{ik} \frac{\partial G_k}{\partial q_j} \tag{3.12.33}$$

$\forall i, j = 1, \ldots, \ell$, with the obvious choice of arguments of these functions; e.g. $\boldsymbol{\pi} = \boldsymbol{\pi}_0, \mathbf{q} = \mathbf{q}_0$. Differentiation of the identity $\boldsymbol{\kappa} \equiv \mathbf{G}(\boldsymbol{\pi}, \mathbf{Q}(\boldsymbol{\pi}, \boldsymbol{\kappa}))$ gives

$$\frac{\partial G_i}{\partial q_j} = (D^{-1})_{ij}, \qquad \frac{\partial G_i}{\partial \pi_j} + \sum_{s=1}^{\ell} \frac{\partial G_i}{\partial q_s} C_{sj} = 0, \tag{3.12.34}$$

$\forall i, j = 1, \ldots, \ell$. More concisely, rewrite Eqs. (3.12.33) and (3.12.34) as

$$\begin{aligned} \frac{\partial \mathbf{F}}{\partial \boldsymbol{\pi}} &= A - BD^{-1}C, \qquad & \frac{\partial \mathbf{F}}{\partial \mathbf{q}} &= BD^{-1}, \\ \frac{\partial \mathbf{G}}{\partial \mathbf{q}} &= D^{-1}, \qquad & \frac{\partial \mathbf{G}}{\partial \boldsymbol{\pi}} &= -D^{-1}C \end{aligned} \tag{3.12.35}$$

and the conditions (3.12.32) become

$$\begin{aligned} &A - BD^{-1}C = (D^{-1})^T, \\ &BD^{-1} = (BD^{-1})^T, \qquad \text{i.e.} \quad BD^{-1} = (D^{-1})^T B^T, \quad \text{i.e } D^T B = BD^T, \\ &D^{-1}C = (D^{-1}C)^T, \qquad \text{i.e.} \quad D^{-1}C = C^T (D^{-1})^T. \quad \text{i.e. } CD^T = DC^T. \end{aligned} \tag{3.12.36}$$

So after checking xthat Eqs. (3.12.36) are a disguised form of the complete canonicity conditions (3.12.11) and the proof will be complete.

In fact, the third of Eqs. (3.12.36) is implied by the second of Eqs. (3.12.11). Furthermore, using the first and the transposition of the third of Eqs. (3.12.11) we see that[23]

$$AB^T = BA^T = BD^{-1}DA^T = BD^{-1}(I + CB^T) = BD^{-1} + B(D^{-1}C)B^T \tag{3.12.37}$$

which shows that BD^{-1} is symmetric since such is $AB^T$and $B(D^{-1}C)B^T$ (having already seen that $D^{-1}C$ is symmetric). So the second of Eqs. (3.12.36) also holds. Finally the first of Eq. (3.12.36) means

$$AD^T - BD^{-1}CD^T = I \tag{3.12.38}$$

which, since $CD^T = DC^T$ by the Eq. (3.12.11), shows that the second equality in Eqs. (3.12.38) simply means that $AD^T - BC^T = I$ which is true, because it is the first of Eqs. (3.12.11). mbe

3.12.1 Problems and Complements

1. Let $\mathcal{C}$ be a map of W onto W' and suppose that there is $\Phi \in C^\infty(G(\mathcal{C}))$, $G(\mathcal{C}) \equiv \{\mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}' \,|\, (\mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}') \in W \times W',\ (\mathbf{p}, \mathbf{q}) = \mathcal{C}(\mathbf{p}', \mathbf{q}')\}$, such that

$$\mathbf{p} \cdot d\mathbf{q} = \mathbf{p}' \cdot d\mathbf{q}' + d\Phi.$$

Show that $\mathcal{C}$ is a time independent completely canonical map and that it is also "action preserving" in the sense that, if λ is a closed curve in W' and $\mathcal{C}\lambda$ is its $\mathcal{C}$-image in W, it is

$$\int_{\mathcal{C}\lambda} \mathbf{p} \cdot d\mathbf{q} = \int_\lambda \mathbf{p}' \cdot d\mathbf{q}'$$

2. Consider in $\mathcal{R}^2$ the annulus $D = \{(q_1, q_2) \,|\, \alpha < q_1^2 + q_2^2 < \beta,\ \alpha, \beta > 0\}$, and let $f(q_1, q_2)dq_1 + g(q_1, q_2)dq_2$ be an exact but non integrable differential form on D. Define

$$\mathcal{C}(p_1, p_2, q_1, q_2) \stackrel{def}{=} (p_1', p_2', q_1', q_2') \equiv (p_1 + f(q_1, q_2), p_2 + g(q_1, q_2), q_1, q_2).$$

Show that it is completely canonical (time independent, of course). (*Hint:* Note that $p_1' dq_1' + p_2' dq_2' = p_1 dq_1 + p_2 dq_2 + f(q_1, q_2)dq_1 + g(q_1, q_2)dq_2$ and recall that every exact form is locally integrable and that the complete canonicity is a local property and use problem (1).)

3. Show that not all completely canonical maps are action preserving in the sense of problem (1). (*Hint:* Consider the map in Problem 2 and choose λ to be the curve $p_1 = p_2 = 0, q_1^2 + q_2^2 = \frac{1}{2}(\alpha + \beta)$.)

4.* Show that the existence of $\Phi \in C^\infty(G(\mathcal{C}))$ verifying the property introduced in Problem (1) is a necessary and sufficient condition in order that $\mathcal{C}$ be an action preserving time independent completely canonical map of W onto W'.

[23] We use the fact that if M is a symmetric $\ell \times \ell$ matrix and E is an arbitrary $\ell \times \ell$ matrix, then $E^T ME$ is a symmetric matrix (exercise).

5. Show that the map $\mathcal{C}(p_x,,p_y,x,y) = (p_1,p_2,q_1,q_2)$ between $\mathcal{R}^2 \times (\mathcal{R} \times \mathcal{R}_+)$/(set of points with $p_x = 0$) and $\mathcal{R}^4$/(set of points with $p_1q_1 + p_2q_2 \le 0$ or with $p_1 \le 0$) defined by the following relation, setting $i = \sqrt{-1}$:

$$p \equiv p_x + i\,p_y = \frac{i}{2}(p_1 + i\,q_2)^2, \qquad q \equiv x + i\,y = \frac{p_2 + i\,q_1}{p_1 + i\,q_2}$$

is completely canonical (time independent). (*Hint:* Check that it is one-to-one and $p_x dx + p_y dy \equiv \mathcal{R}\text{e}\,\overline{p dq} = p_1 dq_1 + p_2 dq_2 - \frac{1}{2}d(p_1q_1 + p_2q_2)$.)

6. Let $(p',q') = \mathcal{C}(p,q)$ be a map from $W \subset \mathcal{R}^2$ onto $W' \subset \mathcal{R}^2$. Show that a necessary and sufficient condition for $\mathcal{C}$ being completely canonical is that it is orientation and area preserving (recall that a map is "orientation preserving" if its Jacobian matrix L has positive determinant). (*Hint:* Note that the matrices A, B, C, D are numbers and therefore (3.12.5) holds if and only if $\det L = 1$.)

7. Extend the notion of completely canonical time independent map by replacing (hence extending) (3.11.52) by $S(\mathbf{m}) = \lambda\Sigma(\boldsymbol{\mu}) + \text{constant}$. Discuss the case $\lambda = -1$ and prove a proposition like Proposition 23. Find the physical meaning of λ.

8.* Consider the Hamiltonian on $\mathcal{R}^2 \times (\mathcal{R} \times \mathcal{R}_+)$: $H_0(p_x,p_y,x,y) = \frac{1}{2}y^2(p_x^2 + p_y^2)$, ("Hamiltonian for the geodesic motion for the geometry $ds^2 = \frac{dx^2+dy^2}{y^2}$ on $\mathcal{R} \times \mathcal{R}_+$, see Problems 19-24, p.230), and show that the canonical map in Problem (5) transforms H_0 into $\frac{1}{8}(p_1q_1 + p_2q_2)^2$. Write and solve the Hamilton's equations in the new coordinates.

9.* In the context of Problem (8) consider the canonical map $p_1' = \frac{p_1+q_1}{\sqrt{2}}, q_1' = \frac{p_1-q_1}{\sqrt{2}}, p_2' = \frac{p_2+q_2}{\sqrt{2}}, q_2' = \frac{p_2-q_2}{\sqrt{2}}$ and show that H is transformed by it into $\frac{1}{2}((p_1')^2 - (q_2')^2 + (p_2')^2 - (q_2')^2)^2$. Interpret this as saying that the geodesic motions of H_0 taking place at a given energy E can be thought of as describing the motions of two independent hyperbolic oscillators (i.e. two particles on a negative quadratic potential). How does this picture change as E varies?

10.* Show that the map $(p_1,p_2,q_1,q_2) \to (p_x,p_y,x,y)$ defined in Problem 5 is one-to-one from $\widetilde{G} = \mathcal{R}^4$/(set of points for which $p_1q_1 + p_2q_2 \le 0$) onto $G' = \mathcal{R}^2 \times \mathcal{R} \times \mathcal{R}_+$/(set of points for which $p_x = p_y = 0$). If however the "opposite" points (p_1,p_2,q_1,q_2) and $(-p_1,-p_2,-q_1,-q_2)$ are identified, the map becomes one-to-one. Then remark that (p_1,p_2,q_1,q_2) may be regarded as coordinates (modulo the sign) for the points of the set $G' = \{\widetilde{G}$ with opposite points identified$\}$ in the same sense as a point $\boldsymbol{\varphi} \in \mathcal{R}^\ell$ can be regarded as a coordinate (modulo 2π) for a point in $\mathcal{T}^\ell$.
Using this remark extend the notion of time independent completely canonical maps to cover the case when W instead of being a subset of $V \times (\mathcal{T}^{\ell_1} \times \mathcal{R}^{\ell_2})$ is a subset of G' and show that the map under consideration is completely canonical, in this new sense, as a map between $\widetilde{G}$ and G'.

11. Try to extend the notion of completely canonical time independent map to maps of arbitrary open surfaces of dimension 2ℓ by abstracting the essential properties of the examples discussed in definition 17, p.220, and in Problem (10) where the 2ℓ-dimensional surfaces are very special, i.e. they are, respectively, of the form $V \times \mathcal{T}^{\ell_1} \times \mathcal{R}^{\ell_2}$, $\ell_1 + \ell_2 = \ell$, $V \subset \mathcal{R}^\ell$, or the set $\widetilde{G}$ with opposite points identified.

12. Let W be the phase for a regular time independent Hamiltonian function H, see observation (4), p.216. Let $T > 0$, $(\mathbf{p},\mathbf{q}) \in W$, and suppose that the solution $S_t(\mathbf{p},\mathbf{q})$ to the Hamiltonian equations with initial datum $(\mathbf{p},\mathbf{q})$ stays in W for all $t \in [0,T]$: $S_t(\mathbf{p},\mathbf{q}) \in W$. Define $F_t(\mathbf{p},\mathbf{q}) \equiv F(S_t(\mathbf{p},\mathbf{q}))$ and show that

$$\frac{dF_t(\mathbf{p},\mathbf{q})}{dt} = \{H, F_t\}(\mathbf{p},\mathbf{q})$$

i.e., "the time derivative of an observable F is given by its Poisson bracket with the Hamiltonian"

13. In the context of (12) show that $\{H, F_t\}(\mathbf{p},\mathbf{q}) \equiv \{H, F\}(S_t(\mathbf{p},\mathbf{q}))$: since in Physics the operation of associating with $F \in C^\infty(W)$ the function $\mathcal{L}F = H, F$ is called the "Liouville's operator action" this can be read: the "Liouville operator commutes with the time evolution".

14. Let E, F, G be in $C^\infty(W)$, where W is the phase space for a regular time independent Hamiltonian H. Show that

$$\{E,\{F,G\}\} + \{F,\{G,E\}\} + \{G,\{E,F\}\} = 0, \qquad \{E,F\} = -\{F,E\}$$
$$\{E, FG\} = \{E,F\}G + \{E,G\}F$$

These relations are called respectively "the Jacobi identity", the "antisymmetry" and the "derivation property" of the Poisson bracket.

15. Show that, in the context of Problem (12), the relations ("Liouville's equations")

$$\frac{dF(S_t(\mathbf{p},\mathbf{q}))}{dt} = \{H, F\}(S_t(\mathbf{p},\mathbf{q})) = \mathcal{L}F(S_t(\mathbf{p},\mathbf{q}))$$

imply, if valid for all $F \in C^\infty(W)$, for all $(\mathbf{p},\mathbf{q}) \in W$ and for t small (depending possibly on $(\mathbf{p},\mathbf{q})$), that $t \to S_t(\mathbf{p},\mathbf{q})$ verifies the Hamilton's equations, ("equivalence between the Hamilton's equations and the Liouville's equations").

Other problems on canonical maps can be found at the end of §4.9-4.12 and §5.10 and §5.12.

16. Let $\mathcal{C}(\boldsymbol{\pi},\boldsymbol{\kappa}) = (\mathbf{p},\mathbf{q})$ be a completely canonical map defined between sets $U, W \subset \mathcal{R}^{2\ell}$. Then the Jacobian determinant of $\mathcal{C}$ is a matrix L with determinant $\det L = 1$. (*Hint:* Write L as $\frac{\partial(\mathbf{p},\mathbf{q})}{\partial\boldsymbol{\pi},\boldsymbol{\kappa}}$ and suppose that $\mathcal{C}$ has a generating function $F(\mathbf{q},\boldsymbol{\kappa})$, *for instance*. Then express $(\mathbf{p},\mathbf{q})$ as functions of $(\boldsymbol{\kappa},\mathbf{q})$ first and remark that the Jacobian of this map is

$$\frac{\partial(\mathbf{p},\mathbf{q})}{\partial\boldsymbol{\kappa},\mathbf{q}} = \begin{pmatrix} -\frac{\partial^2 F}{\partial\boldsymbol{\kappa}\partial\mathbf{q}} & -\frac{\partial^2 F}{\partial\mathbf{q}^2} \\ 0 & 1 \end{pmatrix}$$

whose determinant is $(-1)^\ell \det \frac{\partial^2 F}{\partial\boldsymbol{\kappa}\partial\mathbf{q}}$. Similarly the Jacobian of the map $(\boldsymbol{\kappa},\mathbf{q}) \to (\boldsymbol{\pi},\boldsymbol{\kappa})$ is

$$\frac{\partial(\boldsymbol{\kappa},\mathbf{q})}{\partial(\boldsymbol{\pi},\boldsymbol{\kappa})} = \begin{pmatrix} (\frac{\partial^2 F}{\partial\boldsymbol{\kappa}^2})^{-1} & 1 \\ (\frac{\partial^2 F}{\partial\mathbf{q}\boldsymbol{\kappa}})^{-1} & 0 \end{pmatrix}$$

whose determinant is $(-1)^\ell(\det \frac{\partial^2 F}{\partial\boldsymbol{\kappa}\partial\mathbf{q}})^{-1}$. The identity

$$L = \frac{\partial(\mathbf{p},\mathbf{q})}{\partial(\boldsymbol{\pi},\boldsymbol{\kappa})} = \frac{\partial(\mathbf{p},\mathbf{q})}{\partial(\boldsymbol{\kappa},\mathbf{q})} \cdot \frac{\partial(\boldsymbol{\kappa},\mathbf{q})}{\partial(\boldsymbol{\pi},\boldsymbol{\kappa})}$$

implies, therefore, $\det L = 1$ (from [28] p.199.))

Concluding Comments to Chapter 3

(1) We have described by the word "action" certain quantities which, in fact, do not motivate such a nice name [see Eqs. (3.3.4), etc.]. Actually, in contemporary literature, the convention of calling Eq. (3.3.4) "action of a motion", or "least action principle" the corresponding variational principle, prevails.

This is perhaps historically incorrect: the action was introduced by Maupertuis when he formulated the variational principle bearing his name, Problem (16) p.229.[24] The numerical value of the quantity that Maupertuis called "action of a motion", computed on the real motion developing under the influence of given conservative forces and ideal constraints, is related, in a very simple way, to the value of the action of Eq. (3.3.4) computed on the real motion (see Problem 15, §3.11 p.229). The same occurs for the numerical value of other quantities also sometimes called "action", see, for instance, Eq. (3.11.50). These simple relations explain why there is so much confusion in the names. However, it should be stressed that among the various notions of action there are simple relations only if we compare the numerical values that they have on the real motions: it would not make sense to ask if there is a simple relation between the values taken on the varied motions (mainly because in the different variational principles, the motions are described and parameterized differently and, therefore, one cannot compare them).
(2) It is interesting to quote Maupertuis in connection with his definition of action, afterwards interpreted by Euler as in Problem 16, p.229 (quoted from [31], Chapter III, §2.8):

We must explain what is meant by quantity of action. When a body is moved from one point to another, a certain action is necessary. This action depends upon the velocity of body, upon the space it covers, but it is neither the velocity nor the space separately considered. The greater the body's velocity and the longer the path that it covers, the greater the action; the action is proportional to the sum of the spaces, each multiplied by the speed with which the bodies cover them. It is the quantity of action, the true expenditure of Nature, which she administers with as much economy as possible in the movement of light

The last line refers to Maupertuis' application of his principle to the propagation of light. The other lines are a nice way of saying

$$A = \int_{\boldsymbol{\xi}_1}^{\boldsymbol{\xi}_2} \mathbf{v} \cdot d\mathbf{q} = \frac{1}{m} \int_{\boldsymbol{\xi}_1}^{\boldsymbol{\xi}_2} \mathbf{p} \cdot d\mathbf{q},$$

and the condition of stationarity of A on a motion $t \to (\mathbf{p}(t), \mathbf{q}(t))$ of given energy E can be shown to be equivalent to the stationarity of the quantity in Problem 16, §3.11 (a further problem for the reader).

For a comment on Maupertuis' definition, see the angry pages of E. Mach ([31], Chapter III, §8.4).
(3) To understand the historical development of the various principles, one can consult Mach, where they are critically discussed, paying due attention to history. In his book ([31], Chapter IV, §2), one also finds an interesting com-

[24] The original formulation was, in fact, quite obscure and it was later clarified by Euler (see [31], Ch. III).

ment on the "theological, animistic and mystical points of view in mechanics" (see, also, Observation (2), p.164).
(4) Concrete and interesting exercises for this chapter can be found in the book [32].

For §3.1 and §3.3 see:
Chapter 3 §10, §11, §12;
Chapter 4 §21, §22, §23;
Chapter 9 §26, §27, §28, §29, §30, §31, §32, §33;
Chapter 10 §4, §35, §36, §38.

For §3.3, §3.4, §3.5, and §3.8 see:
Chapter 4, §13, §14;
Chapter 5, §15, §16, §17, §18; Chapter 10, §37, §39, §43;
Chapter 11, §46, §47, §48;
Chapter 6, §19, §20.

For §3.11 and §sec:III-12, see:
Chapter 11 §49.

One can also consult the book [16]
For §3.1 and §3.2 see:
Chapters 6 and 11.
For §3.3, §3.4, §3.5, and §3.8 see:
hapters 2, 3, 5, 7, 8, 10, 12, 13, 14, 17, 18, and 21.

4

Special Mechanical Systems

4.1 Systems of Linear Oscillators

In this chapter we adhere systematically to the convention of denoting and writing the Lagrangian functions that we shall meet as $\mathcal{L}(\dot{\mathbf{x}},\mathbf{x},t)$ or $\mathcal{L}(\dot{\mathbf{x}},\mathbf{x})$ or $\mathcal{L}(\dot{\mathbf{q}},\mathbf{q},t)$, rather than as functions of generic variables $(\boldsymbol{\alpha},\boldsymbol{\beta},t)$: the notation is obviously improper since in such cases the variables $\dot{\mathbf{x}}$ and $\mathbf{x}$ are not Cartesian coordinates but local (or toroidal) coordinates, and often the mechanical systems will be described directly in local coordinates omitting the obvious but tedious discussion necessary when the local coordinates are not global (i.e., they are not globally equivalent to Cartesian coordinates).

A typical example of this situation is when one says that a point mass ideally bound to remain on the sphere of radius ϱ is described by a Lagrangian function given, in polar coordinates, by

$$\mathcal{L}(\dot{\theta},\dot{\varphi},\theta,\varphi,t)=\frac{m}{2}\varrho^2(\dot{\theta}^2+(\sin\theta)^2\,\dot{\varphi}^2) \tag{4.1.1}$$

After a little practice and thought, this notational convention, very common in literature, will appear natural and should not give rise to any confusion.

Hence, a system of linear oscillators, each with 1 degree of freedom, is the mechanical system defined by

$$\mathcal{L}(\dot{\mathbf{x}},\mathbf{x})=\frac{1}{2}\sum_{ij=1}^{\ell}g_{ij}\dot{x}_i\dot{x}_j-\frac{1}{2}\sum_{ij=1}^{\ell}v_{ij}x_ix_j, \tag{4.1.2}$$

where $G = (g_{ij})_{i,j=1,\dots,\ell}$, $V = (v_i j)_{i,j=1,\dots,\ell}$ are two $\ell \times \ell$ symmetric positive-definite matrices (see Appendix F, p.525). The Lagrangian equations corresponding to Eq. (4.1.2) are

$$\sum_{j=1}^{\ell} g_{ij}\ddot{x}_j = -\sum_{j=1}^{\ell} v_{ij}x_j, \qquad i = 1, \dots, \ell. \tag{4.1.3}$$

They can be treated in full generality and their theory is summarized by following proposition stating that essentially Eq. (4.1.3) is equivalent through a "simple" transformation, to ℓ equations of the type:

$$\ddot{y}_i = -\omega_i^2\, y_i, \qquad i = 1, \dots, \ell. \tag{4.1.4}$$

I Proposition. *The most general solution of Eq. (4.1.2) for $t \in \mathcal{R}$ can written in terms of ℓ arbitrary non-negative constants $\mathbf{A} = (A_1, ..., A_\ell)$ and of ℓ angles $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_\ell)$ as*

$$\mathbf{x}(t) = \sum_{i=1}^{\ell} \sqrt{\frac{2A_i}{\omega_i}}\, \boldsymbol{\eta}^{(i)}\, \cos(\omega_i t + \varphi_i), \tag{4.1.5}$$

where $\omega_1, \dots, \omega_\ell$ are the ℓ positive solutions of the ℓ-th order equation for ω^2:

$$\det(-\omega^2 G + V) = 0 \tag{4.1.6}$$

and the vectors $\boldsymbol{\eta}^{(1)}, \dots, \boldsymbol{\eta}^{(\ell)}$ verify the equation:

$$-\omega_i^2\, G\, \boldsymbol{\eta}^{(i)} + V\, \boldsymbol{\eta}^{(i)} = \mathbf{0}, \qquad i = 1, \dots, \ell \tag{4.1.7}$$

and they can be chosen so that

$$(G\, \boldsymbol{\eta}^{(i)}) \cdot \boldsymbol{\eta}^{(j)} = \delta_{ij}, \qquad i, j = 1, \dots, \ell. \tag{4.1.8}$$

Observations.
(1) In Eq. (e4.1.5), one could of course write A_i instead of $\sqrt{2A_i/\omega_i}$: however, the square root is more convenient since in this way the map $(\dot{\mathbf{x}}(0), \mathbf{x}(0)) \to (\mathbf{A}, \boldsymbol{\varphi})$ can be related to a canonical transformation [see Exercises for §4.1 and Observation (3) to Corollary 3, p.249].
(2) Therefore, Eq. (4.1.3) admits periodic solutions like

$$\sqrt{\frac{2A}{\omega}}\, \boldsymbol{\eta}\, \cos(\omega\, t\, +\, \varphi). \tag{4.1.9}$$

Such oscillations are called "normal vibration modes" or "normal motions". The preceding proposition tells us that there exist ℓ (independent) normal modes, orthogonal in the sense of Eq. (4.1.8) and that every oscillation is a "superposition" of normal modes.

To underline the interest of the orthogonality of the normal oscillation modes, let us deduce from Proposition I, and before its proof, the following corollary.

2 Corollary. *The energy of the oscillations in Eq. (4.1.5) is*

$$E = \sum_{i=1}^{\ell} \omega_i\, A_i \tag{4.1.10}$$

i.e. it is the sum of the energies of each normal mode component.

PROOF. The energy is [see §3.11, Observation (1), p.217]

$$E = \frac{1}{2}\sum_{ij=1}^{\ell} g_{ij}\dot{x}_i\dot{x}_j + \frac{1}{2}\sum_{ij=1}^{\ell} v_{ij}x_ix_j, \tag{4.1.11}$$

which can be written in vector form as $E = \frac{1}{2}(G\dot{\mathbf{x}})\cdot\dot{\mathbf{x}}+\frac{1}{2}(V\mathbf{x})\cdot\mathbf{x}$ or, explicitly, from Eq. (4.1.5):

$$\begin{aligned}E =&\frac{1}{2}\sum_{i,j=1}^{\ell}\Big\{\sqrt{\frac{4A_iA_j}{\omega_i\omega_j}}\,\omega_i\omega_j\,\sin(\omega_i\,t+\varphi_i)\sin(\omega_j\,t+\varphi_j)\cdot(\boldsymbol{\eta}^{(i)},G\boldsymbol{\eta}^{(j)})\\ &+\sqrt{\frac{4A_iA_j}{\omega_i\omega_j}}\,\cos(\omega_i\,t+\varphi_i)\cos(\omega_j\,t+\varphi_j)\cdot(\boldsymbol{\eta}^{(i)},V\boldsymbol{\eta}^{(j)})\Big\}\end{aligned} \tag{4.1.12}$$

and, using Eq. (4.1.5), we can replace $(\boldsymbol{\eta}^{(i)},V\boldsymbol{\eta}^{(j)})$ with $\omega_j^2(\boldsymbol{\eta}^{(i)},G\boldsymbol{\eta}^{(j)})$, and by Eq. (4.1.8) plus trigonometry, one realizes that Eq. (4.1.12) becomes Eq. (4.1.10). mbe

PROOF OF PROPOSITION I. Assume the existence of $\omega_1,\ldots,\omega_\ell$, the ℓ positive roots of Eq. (4.1.6), and of ℓ linearly independent vectors $\boldsymbol{\eta}^{(1)},\ldots,\boldsymbol{\eta}^{(\ell)}$ verifying Eq. (4.1.7). Then by direct substitution of Eq. (4.1.5) into Eq. (4.1.3). one sees that the function in Eq. (4.1.5) satisfies, $\forall\mathbf{A}\in\mathcal{R}_+^\ell$, $\forall\boldsymbol{\varphi}=(\varphi_1,\ldots,\varphi_\ell)\in\mathcal{T}^\ell$, the equations Eq. (4.1.3).

It is also easy to see that given $(\boldsymbol{\eta},\boldsymbol{\xi})\in\mathcal{R}^{2\ell}$ arbitrarily, it is possible to determine $\mathbf{A}\in\mathcal{R}_+^\ell$, $\boldsymbol{\varphi}\in\mathcal{T}^\ell$ so that Eq. (4.1.5) verifies the datum $\mathbf{x}(\mathbf{0})=\boldsymbol{\xi}$, $\dot{\mathbf{x}}(0)=\boldsymbol{\eta}$ for $t=0$. In fact the conditions

$$\boldsymbol{\xi}=\sum_{j=1}^{\ell}\sqrt{\frac{2A_j}{\omega_j}}\,\boldsymbol{\eta}^{(j)}\,\cos\varphi_j,\qquad \boldsymbol{\eta}=-\sum_{j=1}^{\ell}\sqrt{2\omega_jA_j}\,\boldsymbol{\eta}^{(j)}\,\sin\varphi_j, \tag{4.1.13}$$

imply, by scalar multiplication of both sides of Eq. (4.1.13) by $G\boldsymbol{\eta}^{(i)}$, $i=1,\ldots,\ell$:

$$(G\boldsymbol{\eta}^{(i)}) \cdot \dot{\boldsymbol{\xi}} = \sqrt{\frac{2A_i}{\omega_i}}\,\boldsymbol{\eta}^{(i)} \cos\varphi_i, \qquad (G\boldsymbol{\eta}^{(i)}) \cdot \boldsymbol{\eta} = -\sqrt{2A_i\omega_i}\,\sin\varphi_i \quad (4.1.14)$$

by Eq. (4.1.8). Equation (4.1.14) determines A_i and φ_i because

$$(\sqrt{A_i}, \varphi_i) = \{\text{polar coordinates of the point with Cartesian coordinates}$$

$$\Big(\sqrt{\frac{\omega_i}{2}}\,G\boldsymbol{\eta}^{(i)} \cdot \boldsymbol{\xi}, \frac{1}{\sqrt{2\omega_i}} G\boldsymbol{\eta}^{(i)} \cdot \boldsymbol{\eta}\Big) \in \mathcal{R}^2 \qquad (4.1.15)$$

Viceversa, if (A_i, φ_i) verifies Eq. (4.1.14), it is easy to see that, since the vectors $\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(\ell)}$ are ℓ linearly independent vectors in $\mathcal{R}^\ell$ (by assumption) and, therefore, form a basis in $\mathcal{R}^\ell$, Eq. (4.1.13) necessarily follows.

By virtue of the existence and uniqueness theorems of differential equations, Eq. (4.1.3) is the most general C^∞ solution to Eq. (4.1.3), $t \in \mathcal{R}$.

It remains to show the actual existence of ℓ linearly independent vectors $\eta^{(1)}, \ldots, \boldsymbol{\eta}^{(\ell)}$ and of ℓ numbers $\omega_1, \ldots, w_\ell > 0$. This is a well-known proposition of algebra (see Appendix F, p.525). mbe

It will be useful to stress a simple corollary of Proposition I. For this purpose, we recall the definition of the ℓ-dimensional torus $\mathcal{T}^\ell$ obtained by identifying opposite sides of the square $[0, 2\pi]^\ell$, see Definitions 12 and 13, p.100 and 101, and that of a function in $C^\infty(\mathcal{T}^\ell)$ and set:

1 Definition. *Given $\boldsymbol{\vartheta} = (\theta_1, \ldots, \theta_\ell) \in \mathcal{R}^\ell$, the transformation of $\mathcal{T}^\ell$ into itself,*

$$\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_\ell) \to \boldsymbol{\varphi} + \boldsymbol{\vartheta} = (\varphi_1 + \theta_1, \ldots, \varphi_\ell + \theta_\ell), \ \bmod 2\pi \qquad (4.1.16)$$

will be called a "rotation of $\mathcal{T}^\ell$ with parameters $\boldsymbol{\vartheta} = (\theta_1, \ldots, \theta_\ell) \in \mathcal{R}^\ell$". The group $(S_t)_{t\in\mathcal{R}}$ of transformations of $\mathcal{T}^\ell$ into itself defined by

$$S_t\boldsymbol{\varphi} = S_t(\varphi_1, \ldots, \varphi_\ell) = (\varphi_1 + \omega_1 t, \ldots, \varphi_\ell + \omega_\ell t), \ \bmod 2\pi() \qquad (4.1.17)$$

will be called the "flow on $\mathcal{T}^\ell$ generated by the rotation of $\mathcal{T}^\ell$ with speed $\boldsymbol{\omega}$ or the "quasi-periodic flow on $\mathcal{T}^\ell$ with pulsation $\boldsymbol{\omega}$".

The following is then a corollary to Proposition I.

3 Corollary. *It is possible to establish a correspondence between all the initial data $(\boldsymbol{\eta}, \boldsymbol{\xi}) \in \mathcal{R}^{2\ell}$ for Eq. (4.1.3) and the set of the points $(\mathbf{A}, \boldsymbol{\varphi}) \in \mathcal{R}_+^\ell \times \mathcal{T}^\ell$ via Eq. (4.1.15).*
The correspondence is one to one, nonsingular, and of class C^∞ between $(0, +\infty) \times \mathcal{T}^\ell$ and its image in $\mathcal{R}^{2\ell}$.
In $(\mathbf{A}, \boldsymbol{\varphi})$ coordinates, the motion of Eq. (4.1.5) is simply

$$t \to (\mathbf{A}, \boldsymbol{\varphi} + \boldsymbol{\omega}\, t), \tag{4.1.18}$$

i.e., it is a quasi-periodic flow on the torus $\{\mathbf{A}\} \times \mathcal{T}^\ell$.

Observations.
(1) Corollary 3 and Eq. (4.1.18) say that the motion of ℓ harmonic oscillators "consists of quasi-periodic motions taking place on a family of ℓ-dimensional tori" parameterized by ℓ parameters $\mathbf{A}$. If one discards the data for which some of the normal modes are at rest (i.e., those for which some of the A's vanish), one can also say that the initial data space can be thought of as "foliated" by an ℓ-dimensional family of ℓ-dimensional tori.
(2) The parameter A_i is called the "action of the i-th normal mode". If one describes the system in $(\mathbf{A}, \boldsymbol{\varphi})$ coordinates in the region where $\mathbf{A} \in (0, +\infty)^\ell$, it is clear that it can be regarded as a Hamiltonian system on $(0, +\infty)^\ell \times \mathcal{T}^\ell$ with Hamiltonian

$$h(\mathbf{A}, \boldsymbol{\varphi}) = \sum_{i=1}^{\ell} \omega_i\, A_i = \boldsymbol{\omega} \cdot \mathbf{A} \tag{4.1.19}$$

which leads immediately to Eq. (4.1.18).
(3) Observation (2) leads us to think that if the original system with Lagrangian (4.1.2) is described in the Hamiltonian form by the Hamiltonian

$$H(\mathbf{p}, \mathbf{x}) = \frac{1}{2} G^{-1}\mathbf{p} \cdot \mathbf{p} + \frac{1}{2} V\mathbf{x} \cdot \mathbf{x} \tag{4.1.20}$$

[see Eq. (3.11.25)], the map $(\mathbf{A}, \boldsymbol{\varphi}) \longleftrightarrow (\mathbf{p}, \mathbf{x})$ between $(0, +\infty)^\ell \times \mathcal{T}^\ell$ and the part of phase space where all the normal modes are excited (i.e. $A_i > 0$, $\forall\, i$) is a completely canonical transformation: this is in fact true and it is the reason for writing Eq. (4.1.5) with $\sqrt{\frac{2A_i}{\omega_i}}$ instead of the simpler A_i (see exercises).

4.1.1 Exercises

1. Using Problems (1), (2), and (33), §3.11, show that the maps $(p, q) \longleftrightarrow (\pi, \kappa)$ with $\pi = \frac{p}{\sqrt{\omega m}}$, $\kappa = q\sqrt{m\omega}$, and $(\pi, \kappa) \longleftrightarrow (\frac{\pi^2 + \kappa^2}{2}, \varphi) \stackrel{def}{=} (A, \varphi)$ with $\varphi = \{$polar angular coordinate of $(\kappa, \pi) \in \mathcal{R}^2\}$ are completely canonical maps. Show that performing such transformations successively, one builds a completely canonical transformation changing $H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}$ into $H = \omega\, A$.

2. Let $H(\mathbf{p}, \mathbf{q}) = \frac{1}{2} G^{-1}\mathbf{p} \cdot \mathbf{p} + \frac{1}{2} V\mathbf{q} \cdot \mathbf{q}$ with G, V being two positive-definite matrices, $\ell \times \ell$. By Problem (33) of §3.11, the map $(\mathbf{p}, \mathbf{q}) \longleftrightarrow (\boldsymbol{\pi}, \boldsymbol{\kappa})$ defined by $\mathbf{p} = \sqrt{G}\,\pi$, $\mathbf{q} = \sqrt{G^{-1}}\boldsymbol{\kappa}$ (see Appendix F, p.525, for the definition and the existence of the positive matrix $\sqrt{G}$ such that $\sqrt{G}^2 = G$) is completely canonical. Show that it transforms H into $\frac{1}{2}\boldsymbol{\pi} \cdot \boldsymbol{\pi} + \frac{1}{2}\widetilde{V}\boldsymbol{\kappa} \cdot \boldsymbol{\kappa}$ with $\widetilde{V} = \sqrt{G^{-1}}V\sqrt{G^{-1}}$.
Let $\mathcal{R}$ be an orthogonal matrix (see Appendix E), transforming $\widetilde{V}$ into a diagonal matrix Ω with diagonal elements $\omega_1^2, \ldots, \omega_\ell^2$, i.e., $R^T\widetilde{V} \cdot R = \Omega$ (see Appendix F for an existence theorem on R). Show that the further completely canonical change of coordinates $\boldsymbol{\pi} = R^T\widehat{\boldsymbol{\pi}}$, $\boldsymbol{\kappa} = (R^T)^{-1}\widehat{\boldsymbol{\kappa}}$ changes H into

$$\frac{1}{2}\widehat{\pi}^2 + \frac{1}{2}\Omega\widehat{\kappa}\cdot\widehat{\kappa} \equiv \frac{1}{2}\sum_{i=1}^{\ell}(\widehat{\pi}_i^2 + \omega_i^2\widehat{\kappa}_i^2)$$

Then, for each i , by further applying the maps in Exercise 1, the Hamiltonian is changed in $\sum_{i=1}^{\ell}\omega_i A_i$: prove also this as well.

3. Check that the variables $(\mathbf{A}, \boldsymbol{\varphi})$ constructed in Exercise 2 are the same as those appearing in Proposition I.

4.2 Irrational Rotations on ℓ-Dimensional Tori

In §4.1 a natural description of the motion of a system of harmonic oscillators was given as a quasi-periodic flow on $\mathcal{T}^\ell$ of the form

$$S_t\boldsymbol{\varphi} = (\boldsymbol{\varphi} + \boldsymbol{\omega}t) = (\varphi_1 + \omega_1 t, \ldots, \varphi_\ell + \omega_\ell t) \tag{4.2.1}$$

Hence it is convenient to analyze a few properties of the quasi-periodic flows.

2 Definition. *The flow of Eq. (4.2.1) is "irrational" if* $(\omega_1, , \ldots, \omega_\ell) \in \mathcal{R}^\ell$ *are "rationally independent" numbers, i.e., if the relation*

$$\mathbf{n}\cdot\boldsymbol{\omega} = \sum_{i=1}^{\ell} n_i\omega_i = 0, \quad n_1, \ldots, n_\ell \text{ integers}, \tag{4.2.2}$$

implies $n_1 = \ldots n_\ell = 0$.

From the definition it follows:

4 Proposition. *Let* $(S_t)_{t\in\mathcal{R}}$ *be a quasi periodic flow defined on* $\mathcal{T}^\ell$ *by Eq. (4.2.1) with* $\boldsymbol{\omega} \in \mathcal{R}^\ell$. *If* $\boldsymbol{\varphi} \in \mathcal{T}^\ell$ *and* $t_0 \in \mathcal{R}$, *the trajectory*

$$\Omega(t_0) = \{\boldsymbol{\varphi}' \,|\, \boldsymbol{\varphi}' = S_t\boldsymbol{\varphi}, \text{ for some } t \geq t_0\} \tag{4.2.3}$$

is dense on $\mathcal{T}^\ell$ *if and only if the flow is irrational.*

Observation. It would be possible to provide a direct proof of Proposition 4 along the lines of the analogous Proposition 27, p.92, §2.20, in the case $\ell = 2$. However, we prefer to give an alternative proof based on the Fourier series and on the following proposition which is interesting in itself.

5 Proposition. *Let* $f \in C^\infty(\mathcal{T}^\ell)$ *and let* $(S_t)_{t\in\mathcal{R}_+}$ *be a flow of the type of Eq. (4.2.1) on* $\mathcal{T}^\ell$ *which is irrational. Then,* $\forall\,\boldsymbol{\varphi} \in \mathcal{T}^\ell$, *the average value*

$$\overline{f}(\boldsymbol{\varphi}) = \lim_{T\to+\infty}\frac{1}{T}\int_0^T f(S_t\boldsymbol{\varphi})dt \tag{4.2.4}$$

exists and is $\boldsymbol{\varphi}$*-independent and equal to*

$$\overline{f} = \int_{\mathcal{T}^\ell} f(\varphi_1', \ldots, \varphi_\ell')\frac{d\varphi_1'\ldots d\varphi_\ell'}{(2\pi)^\ell} \equiv \int_{\mathcal{T}^\ell} f(\boldsymbol{\varphi}')\frac{d\boldsymbol{\varphi}'}{(2\pi)^\ell}. \tag{4.2.5}$$

PROOF. Since $f \in C^\infty(\mathcal{T}^\ell)$, it may be represented as

$$f(\boldsymbol{\varphi}) = \sum_{n_1,\ldots,n_\ell}^{-\infty,+\infty} \widehat{f}_{n_1\ldots n_\ell} e^{i\sum_j n_j\varphi_j} \equiv \sum_{\mathbf{n}\in\mathcal{Z}^\ell} \widehat{f}_{\mathbf{n}}\, e^{i\mathbf{n}\cdot\boldsymbol{\varphi}}, \tag{4.2.6}$$

where $(\widehat{f}_{\mathbf{n}})_{\mathbf{n}\in\mathcal{Z}^\ell}$ are the Fourier harmonics of f (see Proposition 28, p.103), and they decrease faster than any power in in $|\mathbf{n}|$ as $|\mathbf{n}| \to \infty$.

Furthermore, the right-hand side of Eq. (4.2.5) is just $\widehat{f}_{\mathbf{0}}$ [see Eq. (2.21.13]. Then

$$\frac{1}{T}\int_0^T f(S_t\boldsymbol{\varphi})dt = \sum_{\mathbf{n}\in\mathcal{Z}^\ell} \widehat{f}_{\mathbf{n}}\, e^{i\mathbf{n}\cdot\boldsymbol{\varphi}} \Big\{ \frac{1}{T}\int_0^T e^{i\,t\,\mathbf{n}\cdot\boldsymbol{\omega}} \Big\}\, dt \tag{4.2.7}$$

and the series in Eq. (4.2.7) is bounded above by the convergent series

$$\sum_{\mathbf{n}\in\mathcal{Z}^\ell} |\widehat{f}_{\mathbf{n}}| < +\infty \tag{4.2.8}$$

because the number in curly brackets in Eq. (4.2.7) clearly has a modulus not exceeding 1, being an average of numbers of modulus 1. Then we can take the limit in Eq. (4.2.7), as $T \to +\infty$, term by term.

But the integral in the right-hand side of Eq. (4.2.7) is

$$\frac{1}{T}\frac{e^{i\,T\,\mathbf{n}\cdot\boldsymbol{\omega}} - 1}{i\,\mathbf{n}\cdot\boldsymbol{\omega}} \xrightarrow[T\to+\infty]{} 0, \qquad \text{if } \mathbf{n}\cdot\boldsymbol{\omega} \neq 0 \tag{4.2.9}$$

while it is 1 if $\mathbf{n}\cdot\boldsymbol{\omega} = 0$. However, $\mathbf{n}\cdot\boldsymbol{\omega} = 0$ only for $\mathbf{n} = \mathbf{0}$ and all the terms in Eq. (4.2.7) vanish except that with $\mathbf{n} = \mathbf{0}$ as $T \to +\infty$, and Eq. (4.2.5) is proved. mbe

Note that Proposition 5 is also an immediate consequence of Proposition 30, p.105. The same method of proof of Proposition 5 could be used to prove the following proposition which we describe before proving Proposition 4.

6 Proposition. *With the same hypothesis as that of Proposition 5, let $T \in \mathcal{R}, T \neq 0$, and consider the limit*

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{h=0}^{N-1} f(S_{h\,T}\boldsymbol{\varphi}). \tag{4.2.10}$$

Such a limit exists and is given by Eq. (4.2.5) if the $(\ell+1)$ numbers $\omega \stackrel{def}{=} \frac{2\pi}{T}, \omega_1, \ldots, \omega_\ell$ are rationally independent.

Observations.

(1) Proposition 6 is the generalization to the $\ell > 1$ case of the Observations (5) and (6), p.111. The proof is left to the reader as an exercise on the proof

of Proposition 5.
(2) A simple analysis of the proof of the Propositions 5 and 6 allows us to conclude that the limits of Eqs. (4.2.4) and (4.2.10) exist in general, but they will not generally be $\boldsymbol{\varphi}$ independent unless $\omega_1, \ldots, \omega_\ell$ (or $\frac{2\pi}{T}, \omega_1, \ldots, \omega_\ell$) are rationally independent.

An immediate corollary to Proposition 5 is the following proof of Proposition 4.

Proof of proposition 4. Assume that S_t is an irrational flow. Let $\boldsymbol{\varphi}_0 \in \mathcal{T}^\ell$ and let $\chi \in C^\infty(\mathcal{T}^\ell)$ be a non-negative function having the value 1 in $\boldsymbol{\varphi}_0$, and zero outside a small ball $\sigma_\varepsilon \subset \mathcal{T}^\ell$ with center $\boldsymbol{\varphi}_0$ and radius ε in the metric of $\mathcal{T}^\ell$ [see Eq. (2.21.5), p.101.]

Apply Proposition 5 to χ. We see that the average value of $t \to \chi(S_t\boldsymbol{\varphi})$ cannot approach zero, $\forall\, \boldsymbol{\varphi} \in \mathcal{T}^\ell$. Hence, for every t_0, there must be $t > t_0$ such that $\chi(S_t\boldsymbol{\varphi}) > 0$, i.e., $S_t\boldsymbol{\varphi}$ is closer to $\boldsymbol{\varphi}_0$ than ε. This means that $\Omega(t_0)$ is dense. Viceversa, if there exist integers $\overline{n}_1, \ldots, \overline{n}_\ell$ not all equal to zero such that $\overline{\mathbf{n}} \cdot \boldsymbol{\omega} = 0$, the function on $\mathcal{T}^\ell$ defined by

$$\boldsymbol{\varphi} \to \cos(\overline{\mathbf{n}} \cdot \boldsymbol{\varphi}) \tag{4.2.11}$$

is not constant on $\mathcal{T}^\ell$ but is constant on the trajectory $t \to S_t(\boldsymbol{\varphi})$, $t \in \mathcal{R}_+$, for all $\boldsymbol{\varphi} \in \mathcal{T}^\ell$ (since $\overline{\mathbf{n}} \cdot \boldsymbol{\omega} = 0$). Therefore, for instance, the origin trajectory of the origin cannot approach too closely any point $\boldsymbol{\varphi}$ such that $\cos \boldsymbol{\varphi} \cdot \overline{\mathbf{n}} < 1$ and vice versa. So $\Omega(t_0)$ is not dense. mbe

In the same way in which Proposition 5 implies Proposition 4, one sees that Proposition 6 implies the following corollary.

7 Corollary. *With the same hypothesis as that of Proposition 4, let $\tau > 0$. The denumerable subset of $\mathcal{T}^\ell$,*

$$\Omega_\tau(t_0) = \{\boldsymbol{\psi} \,|\, \exists\, h \text{ integer } h\tau \geq t_0, \boldsymbol{\psi} = S_{h\,\tau}\boldsymbol{\varphi}\,\} \tag{4.2.12}$$

is dense in $\mathcal{T}^\ell$ if and only if the $\ell + 1$ numbers $\omega, \omega_1, \ldots, \omega_\ell$, $\omega = 2\pi\tau^{-1}$, are rationally independent.

Proof. Exercise.

4.3 Ordered Systems of Oscillators. Phenomenological Discussion and Heuristic Formulation of the Model of the Perfect Elastic Body (String, Film, and Solid)

In applications, serious difficulties may be met in the use of the general theory of §4.1, and §4.2. Such use, in fact, presupposes the actual possibility of constructing the proper pulsations $\omega_1, \ldots, \omega_\ell$ and the respective eigenvectors $\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(\ell)}$: their construction, in fact, passes through the solution of an ℓ-th

degree algebraic equation, Eq. (4.1.6), and of ℓ linear systems of ℓ equations, Eq. (4.1.7).

However, it is also true that in important applications, the matrices G and V of §4.1 are not arbitrary, but rather they have special properties sometimes permitting the explicit solution of the normal modes construction.

In §4.3-4.6, some of the most interesting cases will be examined, while this section is devoted to the precise mathematical formulation of the models that will be considered.

Let $\mathcal{Z}_a^d$ be the d-dimensional lattice of the points $\boldsymbol{\xi} \in \mathcal{R}^d$ with coordinate which are integer multiples of $a > 0$:

$$\boldsymbol{\xi} = (n_1 a,\, n_2 a, \ldots, n_d a), \qquad n_1, \ldots, n_d \,\text{integers} \tag{4.3.1}$$

Imagine that around every site $\boldsymbol{\xi} \in \mathcal{Z}_a^d$, a mass m oscillates bound by ideal constraints to move on a straight line through $\boldsymbol{\xi}$ and orthogonal to $\mathcal{R}^d$.

Furthermore, suppose that if $y_{\boldsymbol{\xi}}$ is the elongation with respect to $\boldsymbol{\xi}$ of the oscillator in $\boldsymbol{\xi}$ then:

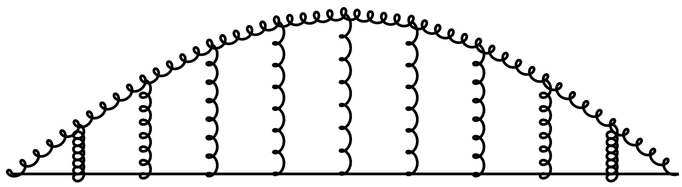


Figure 4.1: chain of oscillators elastically bound by nearest neighbors and to centers aligned on an axis orthogonal to the vibrations.

(i) Every oscillator is subject to a restoring elastic force with potential energy

$$\frac{K}{2}\, y_{\boldsymbol{\xi}}^2, \tag{4.3.2}$$

(ii) Every oscillator is subject to an external force with potential energy

$$m\, g(\boldsymbol{\xi})\, y_{\boldsymbol{\xi}}, \tag{4.3.3}$$

where $g \in C^\infty(\mathcal{R}^d)$ ("weight").
(iii) Between the oscillators adjacent in $\mathcal{Z}_a^d$, an elastic force acts whose potential energy is

$$\frac{1}{2} K' \,[(y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}'})^2 + a^2], \tag{4.3.4}$$

where $|\boldsymbol{\xi}' - \boldsymbol{\xi}| = a$ and the term in square brackets represents the square of the elongation of a spring between the two oscillators.
(iv) An ideal constraint forcing all the oscillators outside an open connected bounded region Ω, with boundary $\partial\Omega$ which is a C^∞-regular surface, to have zero elongation. Set $\Omega_a = \Omega \cap \mathcal{Z}_a^d$.

Only consider the cases $d = 1$ or $d = 2$ will be considered. The $d = 3$ case being a not too interesting model of an elastic solid since it can only "vibrate

in one direction". The situation in the $d = 1$ case is pictured in Fig. 4.1 while the $d = 2$ case is pictured in Fig. 4.2.

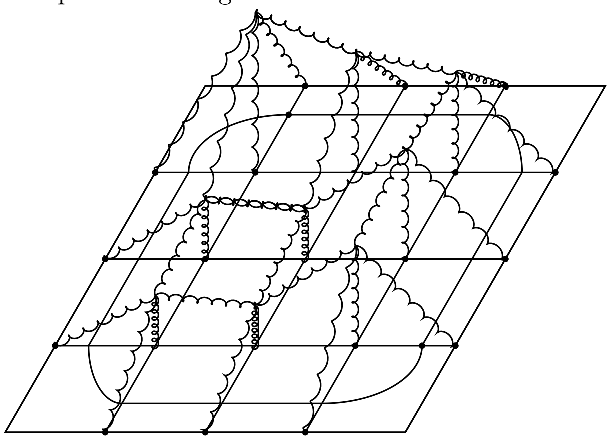


Figure 4.2: System of oscillators elastically bound to their nearest neighbors and to a lattice of centers on a plane orthogonal to the vibrations.

Analytically, the system is described by the Lagrangian function:

$$\begin{aligned}\mathcal{L}_0 =& \frac{1}{2} \sum_{\boldsymbol{\xi}\in\Omega_a} m\,\dot{y}_{\boldsymbol{\xi}}^2 - \frac{K}{2} \sum_{\boldsymbol{\xi}\in\Omega_a} y_{\boldsymbol{\xi}}^2 - m \sum_{\boldsymbol{\xi}\in\Omega_a} g(\boldsymbol{\xi}) y_{\boldsymbol{\xi}} \\ &- \frac{K'}{2} \sum_{\boldsymbol{\xi}\in\Omega_a} \sum_{\mathbf{e}} \frac{1}{\nu(\mathbf{e},\boldsymbol{\xi})} (y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}+a\,\mathbf{e}})2,\end{aligned} \tag{4.3.5}$$

where $\Sigma_{\mathbf{e}}$ denotes the sum over the $2d$ unit vectors directed as the axes of $\mathcal{Z}_a^d$: $\mathbf{e} = \mathbf{e}_1, -\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_2, \ldots, \mathbf{e}_d, -\mathbf{e}_d$ are the d unit vectors associated with $\mathcal{Z}^d$, and, to avoid double counting, $\nu(\mathbf{e}, \boldsymbol{\xi}) = 2$ if $\boldsymbol{\xi}, \boldsymbol{\xi} + a\,\mathbf{e} \in \Omega_a$, $\nu(\mathbf{e}, \boldsymbol{\xi}) = 1$ otherwise.

In the last sum in the right-hand side of Eq. (4.3.5), the term a^2 appearing in Eq. (4.3.4) has been dropped since it produces an additive constant to $\mathcal{L}_0$ (dynamically irrelevant).

In Eq. (4.3.5) there appear terms $y_{\boldsymbol{\xi}}$, with $\boldsymbol{\xi} \notin \Omega_a$ (in fact, if $\boldsymbol{\xi}$ is close to $\partial\Omega$ it can happen that $\boldsymbol{\xi} + a\,\mathbf{e} \notin \Omega_a$. Such terms, conforming to (iv), must be interpreted by setting $y_{\boldsymbol{\xi}} = 0$.

From a physical viewpoint, the interest of the mechanical system in Eq. (4.3.5) lies in the fact, suggested by the above pictures, that if a is very small, it can be considered as a discrete model for an elastic string or film (if $d = 1$ or $d = 2$).

We can imagine that for small a, every "regular" initial datum $(\dot{y}_{\boldsymbol{\xi}}, y_{\boldsymbol{\xi}})_{\boldsymbol{\xi}\in\Omega_a}$, i.e., every datum having the form

$$\dot{y}_{\boldsymbol{\xi}} = u(\boldsymbol{\xi}),\ \dot{y}_{\boldsymbol{\xi}} = v(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in \Omega_a \tag{4.3.6}$$

where u, v are functions in $C^\infty(\mathcal{R}^d)$ vanishing outside Ω, a space that will be denoted $C_0^\infty(\Omega)$, evolves remaining approximately regular, thus simulating the motion of a string or film. In order for this to occur, it is, however, clear that the parameters m, K, K' must be suitably chosen as functions of a: their choice, which we adopt in the following, is motivated by a heuristic discussion.

(a) The mass m of each oscillator must have the form

$$m = \mu\, a^d, \qquad \mu > 0, \tag{4.3.7}$$

since each oscillator should intuitively correspond to a small piece of the body with dimension a: the body will then have density μ.
(b) The constants K, K' have to be determined so as to produce forces proportional to a^d on the oscillator in $\boldsymbol{\xi}$; otherwise their effects would vanish in the $a \to 0$ limit (if $\ll a^d$) or they would produce infinite accelerations (if $\gg a^d$). Hence, since the force associated with K is $-Ky$, it must be:

$$K = \sigma\, a^d, \qquad \sigma > 0 \tag{4.3.8}$$

The force exerted by the two oscillators in $\boldsymbol{\xi} - a\,\mathbf{e}$ and $\boldsymbol{\xi} + a\,\mathbf{e}$ on the oscillator in $\boldsymbol{\xi}$ is

$$-K'[(y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}+a\,\mathbf{e}_i}) + (y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}-a\mathbf{e}_i})], \tag{4.3.9}$$

and if $y_{\boldsymbol{\xi}}$ can be assimilated to $u(\boldsymbol{\xi})$, $u \in C_0^\infty(\Omega)$, we can compute Eq. (4.3.9) using the Taylor-Lagrange expansion to second order as

$$y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}\pm a\mathbf{e}_i} \simeq u(\boldsymbol{\xi}) - u(\boldsymbol{\xi} \pm a\,\mathbf{e}_i) = \mp a\,\partial_i u(\boldsymbol{\xi}) - \frac{a^2}{2}\,\partial_i^2 u(\boldsymbol{\xi}) + O(a^3), \tag{4.3.10}$$

where $\partial_i u, \partial_i^2 u$ are short notations for $\frac{\partial u}{\partial \xi_i}, \frac{\partial^2 u}{\partial \xi_i^2}$. Then Eq. (4.3.9) becomes

$$K' a^2\, \partial_i^2 u(\boldsymbol{\xi}) + O(a^3) \tag{4.3.11}$$

which indicates that it must be set

$$K' a^2 = \tau\, a^d, \qquad \tau > 0. \tag{4.3.12}$$

With the above choices of K, m, K', Eq. (4.3.5) becomes

$$\begin{aligned}\mathcal{L}_0^{(a)} =& \frac{\mu}{2} a^d \sum_{\boldsymbol{\xi}\in\Omega_a} \dot{y}_{\boldsymbol{\xi}}^2 - \frac{\sigma}{2} a^d - \frac{\sigma}{2} a^d \sum_{\boldsymbol{\xi}\in\Omega_a} y_{\boldsymbol{\xi}}^2 - \mu a^d \sum_{\boldsymbol{\xi}\in\Omega_a} g(\boldsymbol{\xi})\, y_{\boldsymbol{\xi}} \\ &- \frac{\tau}{2} a^d \sum_{\boldsymbol{\xi}\in\Omega_a} \sum_{\mathbf{e}} \frac{1}{\nu(\mathbf{e}, \boldsymbol{\xi})} \frac{(y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}+a\,\mathbf{e}})^2}{a^2}.\end{aligned} \tag{4.3.13}$$

This model is not yet completely correct from a physical point of view. The heuristic discussion so far presented has been dealt with by supposing that $\boldsymbol{\xi}$

was far from $\partial\Omega$: if $\boldsymbol{\xi}$ is adjacent to $\partial\Omega$, it is not quite clear what is meant by $y_{\boldsymbol{\xi}}$ being regular since the functions u, v approximating it in Eq. (4.3.6) cannot be a independent, as supposed. A look at Fig. 4.3 suffices to realize this. The points outside Ω and adjacent to it have a rather erratic structure and, quite delicately, are a dependent.

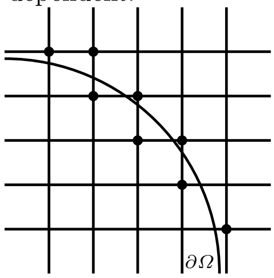


Figure 4.3: The erratic mismatches between the regular lattice and the boundary of Ω.

Though this point may superficially appear irrelevant, it in fact has some importance at least as far as the correct formulation of the meaning of "regular datum $y_{\boldsymbol{\xi}}, \dot{y}_{\boldsymbol{\xi}}$" is concerned.

In the $d = 1$ case, the difficulty can be simply avoided by supposing that a is chosen always so that $\partial\Omega$ (which now consists of two points) is always on $\mathcal{Z}^1_a$: in this case, therefore, we shall actually do so and we shall assume that the system (4.3.13), with the above restriction on the "allowed values" of a, is a "vibrating" or "elastic string" model.

In the $d = 2$ case, it is obviously not possible to circumvent so easily the difficulty and, to understand what to do: let us again refer to some heuristic physical considerations.

When one imagines an elastic homogeneous film oscillating with a fixed boundary $\partial\Omega$, one probably has in mind the following situation: one deposits an elastic homogeneous film on a plane and then "glues" the film on the plane at $\partial\Omega$ and, *afterwards*, lets it oscillate and studies (or watches) the oscillations.

When the surface is described, as in our case, by linked oscillators, the corresponding procedure is that of setting the oscillators in their equilibrium positions on $\mathcal{Z}^a$ and then pinching (with "glue" or "nails") the springs connecting the points $\boldsymbol{\xi} \in \Omega$ to the points $\boldsymbol{\xi}' = \boldsymbol{\xi} + a\,\mathbf{e} \notin \Omega$ at the point $\boldsymbol{\xi} + \varepsilon\,\mathbf{e}$ where the segment $\overline{\boldsymbol{\xi}\boldsymbol{\xi}'}$ crosses $\partial\Omega$. Once this is done, the system is allowed to oscillate.

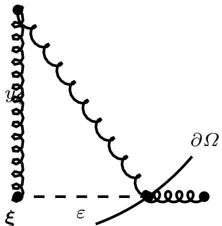


Figure 4.4: The pinching to adapt to the boundary condition.

On the boundary of Ω, the situation drawn in Fig. 4.4. is produced. This means that the elastic constant binding $y_{\boldsymbol{\xi}}$ to $\partial\Omega$ is different from K' contrary to what, instead, is hypothesized in $\mathcal{L}_0^{(a)}$, Eq. (4.3.13). In fact, $y_{\boldsymbol{\xi}}$ is pulled from $\partial\Omega$ by a spring with elastic constant

$$\widetilde{K} = K'\frac{a}{\varepsilon} \tag{4.3.14}$$

because the elastic constant of a piece of spring with elastic constant K' obtained by pinching it at a distance ε when the spring is elongated by a is given by Eq. (4.3.14).

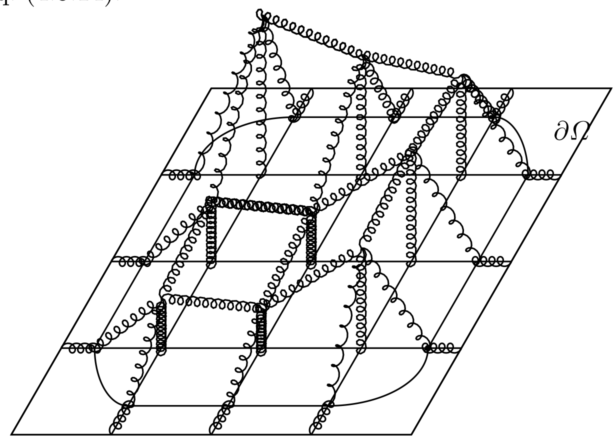


Figure 4.5: Illustration of the system of oscillators corresponding to Eq. (4.3.16).

Then, for $\boldsymbol{\xi} \in \Omega_a$, we set

$$\begin{aligned} \varepsilon_d(\boldsymbol{\xi}, a) =& a \qquad \text{if} \quad \boldsymbol{\xi} + a\,\mathbf{e} \in \Omega_a \\ \varepsilon_d(\boldsymbol{\xi}, e) =& \{\text{distance between } \boldsymbol{\xi} \text{ and } \partial\Omega \cap \overline{\boldsymbol{\xi}(\boldsymbol{\xi} + a\,\mathbf{e})} \quad \text{otherwise} \end{aligned} \tag{4.3.15}$$

and the above considerations are summarized in the following Lagrangian function which will be supposed to be our discrete model of the elastic string or film (see Fig. 4.5), discarding the simpler but more naive model of Eq. (4.3.13):

$$\begin{aligned} \mathcal{L}^{(a)} =& \frac{\mu}{2} a^d \sum_{\boldsymbol{\xi} \in \Omega_a} \dot{y}_{\boldsymbol{\xi}}^2 - \frac{\sigma}{2} a^d - \frac{\sigma}{2} a^d \sum_{\boldsymbol{\xi} \in \Omega_a} y_{\boldsymbol{\xi}}^2 - \mu a^d \sum_{\boldsymbol{\xi} \in \Omega_a} g(\boldsymbol{\xi})\, y_{\boldsymbol{\xi}} \\ & - \frac{\tau}{2} a^d \sum_{\boldsymbol{\xi} \in \Omega_a} \sum_{\mathbf{e}} \frac{1}{\nu(\mathbf{e}, \boldsymbol{\xi})} \frac{a}{\varepsilon_a(\boldsymbol{\xi}, \mathbf{e})} \frac{(y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}+a\,\mathbf{e}})^2}{a^2}. \end{aligned} \tag{4.3.16}$$

Here the values of $y_{\boldsymbol{\xi}'}$ when $\boldsymbol{\xi}' \notin \Omega$, present in Eq. (4.3.16) if $\boldsymbol{\xi}$ is close to $\partial\Omega$ and $\boldsymbol{\xi} + \varepsilon_a(\boldsymbol{\xi}, \mathbf{e}) = \boldsymbol{\xi}' \in \partial\Omega$, have to be thought of as vanishing. Or, more

generally, we may fix the film or string at preassigned elongations on $\partial\Omega$, described by a function $h \in C^\infty(\partial\Omega)$.[1]

In this case, the values of $y_{\boldsymbol{\xi}'}$, for the above $\boldsymbol{\xi}$'s are to be thought of as given by

$$y_{\boldsymbol{\xi}'} = h(\boldsymbol{\xi}') \tag{4.3.17}$$

It is clear that Eq. (4.3.16) differs from Eq. (4.3.13) only because of the terms for which $\boldsymbol{\xi}$ is adjacent to $\partial\Omega$.

It is also clear that the critique to Eq. (4.3.13) raised above can no longer be applied. For instance, if $h \equiv 0$, the initial datum $y_{\boldsymbol{\xi}}, y_{\boldsymbol{\xi}}$, very naturally, can be called "regular" if, $\forall\, \boldsymbol{\xi} \in \Omega_a$,

$$y_{\boldsymbol{\xi}} = u(\boldsymbol{\xi}), \quad \dot{y}_{\boldsymbol{\xi}} = v(\boldsymbol{\xi}) \tag{4.3.18}$$

and $u, v \in C_0^\infty(\overline{\Omega}) = \{$ set of the C^∞ functions defined in a neighborhood of Ω and vanishing outside $\overline{\Omega}\}$.

In the upcoming sections, we shall study some properties of the motions of the system in Eq. (4.3.16) and (4.3.17), paying attention to the problem of regularity for the motions with initial conditions Eq. (4.3.18) and to their interpretability as motions of a string or film.

If $d = 3$, Eq. (4.3.16) still makes sense, but it not longer provides a natural model of an elastic solid. However, it becomes much more natural if $y_{\boldsymbol{\xi}}$, instead of being a scalar quantity ($y_{\boldsymbol{\xi}} \in \mathcal{R}$), is thought of as a vector in $\mathcal{R}^3$. In this case, by thinking $y_{\boldsymbol{\xi}} \in \mathcal{R}^3$, instead of $y_{\boldsymbol{\xi}} \in \mathcal{R}$ (as done so far), Eq. (4.3.16) would yield an interesting (though rather special) model for the elastic deformations of a solid. However, the case $d = 3$ will not be further examined.

4.4 Oscillator Chains and the Vibrating String

Consider the Lagrangian function of Eqs. (4.3.16) and (4.3.17), supposing $\Omega = [0, L]$ and a such that $L/a = N$ is an integer.

Therefore, this function describes a system of $N + 1$ oscillators, the first and the last of which are fixed at given heights. The Lagrangian of Eqs. Eq. (4.3.16) and (4.3.17) becomes

$$\sum_{i-1}^{N-1} \Big(\frac{\mu}{2} a\dot{y}_{ia}^2 + \mu a g(ia) i y_{ia}^2 - \frac{\sigma}{2} a\, y_{ia}^2\Big) - \frac{\tau}{2} a \sum_{i=0}^{N} \frac{(y_{ia} - y_{ia+a})^2}{a^2}, \tag{4.4.1}$$

$$y_0 = h_0, \quad y_L = h_L, \quad g \in C^\infty(R). \tag{4.4.2}$$

[1] A function f defined on a regular surface $\Sigma \subset \mathcal{R}^d$ is said to be in $C^\infty(\Sigma)$ if in any local system $(U, \boldsymbol{\Xi})$ of regular coordinates, its restriction to $\Sigma \cap U$ is a C^∞ function of the coordinates of the points of $U \cap \S$ in $(U, \boldsymbol{\Xi})$ (see Definition 10, §3.6, p.170).

The equations of motion for Eqs. (4.4.1) and (4.4.2) become

$$\mu \ddot{y}_{ia} = \mu a g(ia) - \sigma a y_{ia} - \frac{\tau}{a}(2y_{ia} - y_{ia+a} - y_{ia-a}) \tag{4.4.3}$$

for $i = 1, \ldots, N-1$, where $y_0 = h_0$, $y_L = h_L$, $\mu > 0, \sigma \geq 0, \tau > 0$.

This is a system of linear non homogeneous differential equations which, as usual, we shall study by writing its solutions as sums of a particular solution and of a solution of the homogeneous equation, which is obtained by setting $g = 0$ and $h_0 = h_L = 0$.

Let us first study the homogeneous equation. The results of the following analysis are summarized by Proposition 9 at the end of this section.

In the homogeneous case, Eq. (4.4.3) correspond to the Lagrangian equations Eq. (4.4.1) and (4.4.2) with $g = 0, h_0 = h_L = 0$. This is a system of oscillators of the type considered in §4.1 with, $i, j = 1, \ldots, N-1$,

$$G_{ij} = \mu\, a\, \delta_{ij}, \tag{4.4.4}$$

$$V_{ij} = \sigma\, a\, \delta_{ij} + \frac{\tau}{a}(2\delta_{ij} - \delta_{ij+1} - \delta_{ij-1}), \tag{4.4.5}$$

This can be checked immediately by noting that if $\boldsymbol{\gamma} = (\gamma_i)_{i=1,\ldots,N-1}$, one finds (setting $\gamma_0 = \gamma_N = 0$) that Eq. (4.4.5) yields

$$\sum_{i,j=1}^{N-1} V_{ij}\gamma_i\gamma_j = a\sigma \sum_{j=1}^{N-1} \gamma_j^2 + \frac{\tau}{a} \sum_{j=0}^{N-1} (\gamma_i - \gamma_{i+1})^2 \tag{4.4.6}$$

To solve the system $\omega^2 \boldsymbol{\eta} - V\boldsymbol{\eta} = \mathbf{0}$ [see Eqs. (4.1.6) and (4.1.7)] remark that such a system has the explicit form

$$-\mu\omega^2 \eta_{ja} = -\sigma\eta_{ja} - \tau\frac{(2\eta_{ja} - \eta_{ja+a} - \eta_{ja-a})}{a^2}, \tag{4.4.7}$$

where $j = 1, \ldots, N-1$ and $\eta_0 = \eta_L = 0$.

The manifest analogy between this equation and the linear differential equation $-\omega^2\eta = -\sigma\eta - \tau\eta''$, suggests to look for solutions of Eq. (4.4.7) having the form

$$\eta_{ja} = \sum_{\varrho} \beta_\varrho e^{\alpha_\varrho\, ja}, \qquad \beta_\varrho, \alpha_\varrho \in \mathcal{C}, \tag{4.4.8}$$

where ϱ is a summation index.

In order that $e^{\alpha_\varrho ja}$ is a solution of Eq. (4.4.7) for $j = 2, \ldots, N-2$, it must be [by substitution of Eq. (4.4.8) into Eq. (4.4.7), $j = 2, \ldots, N-2$]:

$$(-\mu\omega^2 + \sigma) = -\frac{2\tau}{a^2}\big(1 - \frac{e^{\alpha_\varrho a} + e^{-\alpha_\varrho a}}{2}\big) \tag{4.4.9}$$

If ω^2 is such that this equation for α_ϱ has a solution α, then $-\alpha$ is also a solution. Hence it seems natural to try to solve Eq. (4.4.7) with $\boldsymbol{\eta}$ given by

$$\eta_{ja} = \beta_+ \, e^{\alpha\, ja} + \beta_- \, e^{-\alpha\, ja}, \qquad j = 1, \dots, N-1, \tag{4.4.10}$$

where α and ω are related by Eq. (4.4.9).

The only equations of the system of Eq. (4.4.7) that Eq. (4.4.10) still may fail to verify are the first and last. If $\boldsymbol{\eta}$ has the form of Eq. (4.4.10), such equations become equations for $\beta_\pm$:

$$\sum_{\varrho=\pm} \big((-\mu\omega^2 + \sigma) + \frac{\tau}{a^2}(2 - e^{-\varrho\,\alpha\, a})\big)\beta_\varrho e^{\varrho\,\alpha\,(N-1)} a = 0 \tag{4.4.11}$$

corresponding to Eq. (4.4.7) with $i = (N-1)$ or, for $i = 1$,

$$\sum_{\varrho=\pm} \big((-\mu\omega^2 + \sigma) + \frac{\tau}{a^2}(2 - e^{-\varrho\,\alpha\, a})\big)\beta_\varrho e^{\varrho\,\alpha\, a} = 0 \tag{4.4.12}$$

which, by using Eq. (4.4.9), become, respectively,

$$\sum_{\varrho=\pm} \frac{\tau}{a^2} e^{\varrho\,\alpha\, Na} \beta_\varrho = 0, \qquad \sum_{\varrho=\pm} \frac{\tau}{a^2} \beta_\varrho = 0. \tag{4.4.13}$$

The latter two homogeneous equations, in the two unknowns β_+ and β_-, have a nontrivial solution if the determinant of the coefficients vanishes, i.e., it must be

$$e^{2\alpha\, Na} = 1 \tag{4.4.14}$$

and, in this case, $\beta_+ = -\beta_-$. Hence, ($i = \sqrt{-1}$):

$$\alpha = i\,\frac{\pi}{Na}\,h, \qquad h = O, 1, \dots, N-1, \dots \tag{4.4.15}$$

to which correspond the solutions [see Eq. (4.4.10)]

$$\eta_{ja}^{(h)} = \beta \, \sin \frac{\pi}{N} h\, j, \qquad h = 0, 1, \dots, N-1 \tag{4.4.16}$$

with the respective eigenvalues ω_h^2 given by Eq. (4.4.9):

$$\omega_h^2 = \frac{\sigma}{\mu} + \frac{\tau}{\mu} \frac{2(1 - \cos \frac{\pi\, h}{Na} a)}{a^2}. \tag{4.4.17}$$

The $N-1$ solutions (4.4.16) are linearly independent vectors: they are, in fact, orthogonal. This follows from the general theory of Appendix F, p.525, since $\omega_{h+1}^2 > \omega_h^2$, $h = 1, \dots, N-2$, but the following direct check is somewhat instructive. Let, in fact, $1 \le h, h' \le N-1$; then[2]

[2] Since $\cos\varphi = \mathcal{R}\mathrm{e}\,(e^{i\varphi})$.

$$\begin{aligned}\sum_{j=1}^{N-1} \eta_{ja}^{(h)} \eta_{ja}^{(h')} \equiv & \beta^2 \sum_{j=1}^{N-1} \sin \frac{\pi\, h}{N} j \, \sin \frac{\pi\, h'}{N} j \\ = & \frac{\beta^2}{2} \sum_{j=1}^{N-1} \Big(\cos \frac{\pi(h-h')j}{N} - \cos \frac{\pi(h+h')j}{N} \Big) \\ = & \frac{\beta^2}{2} \sum_{j=1}^{N-1} \mathcal{R}\mathrm{e} \left(e^{i\pi(h-h')j/2} - e^{i\pi(h+h')j/2} \right) \\ = & \frac{\beta^2}{2} \mathcal{R}\mathrm{e} \Big(\frac{e^{i\pi(h-h')}-1}{e^{i\pi(h-h')/N}-1} - \frac{e^{i\pi(h+h')}-1}{e^{i\pi(h+h')/N}-1} \Big)\end{aligned}$$

which, if $h = h'$, has to be interpreted as $\beta^2 N/2$ and, if $h \neq h'$, is zero since $e^{i\pi(h-h')} = e^{i\pi(h+h')} = \pm 1$ and $\mathcal{R}\mathrm{e}\,(e^{i\alpha}-1) \equiv \frac{1}{2} \equiv -\frac{1}{2}$, $\forall\, \alpha \in \mathcal{R}$. Therefore,

$$\boldsymbol{\eta}^{(h)} \cdot \boldsymbol{\eta}^{(h')} = \beta^2 \frac{N}{2} \delta_{h'}. \tag{4.4.18}$$

Hence, using the results of §4.1, the most general motion of the $N-1$ oscillators described by Eqs. (4.4.1) and (4.4.2) with $h_0 = h_L = 0$ and $g = 0$ is, $\forall\, j = 1, \ldots, N-1$,

$$y_{ja} = \sum_{h=1}^{N-1} A_h \sqrt{\frac{2}{N}}\, (\sin \frac{jh}{Na}\, ja)\, \cos(\omega_h t + \varphi_h). \qquad h = 1, \ldots, N-1, \tag{4.4.19}$$

where $\omega_h > 0$ is given by Eq. (4.4.17) and $A_h \geq 0$, $\varphi_h \in [0, 2\pi]$ are arbitrary constants.

A particular solution to Eq. (4.4.3) can be found as follows. Obviously, the simplest particular solution is, if existing, a stationary one, $\overline{\mathbf{y}}(t) = \mathbf{c}$, i.e. a solution of the system

$$\sigma\, c_{ja} + \tau \frac{2c_{ja} - c_{ja+a} - c_{ja-a}}{a^2} = \mu g(ja) + \frac{\tau}{a^2} (\delta_{j,N-1} h_L + \delta_{j,1} h_0) \tag{4.4.20}$$

for $j = 1, \ldots, N-1$, where $c_0 = c_L = 0$. These equations immediately follow from Eq. (4.4.3) in which the terms with the time derivatives have been eliminated and the inhomogeneous terms depending on g and h have been brought to the right-hand side.

Call $\boldsymbol{\gamma}$ the vector $\boldsymbol{\gamma} = (\gamma_{ia})_{i=1,\ldots,N-1}$ defined by the right-hand side of Eq. (4.4.20). Recalling the definition of V, Eq. (4.4.5), Eq. (4.4.20) can be written as

$$a^{-1}\, V\, \mathbf{c} = \boldsymbol{\gamma}. \tag{4.4.21}$$

This equation has one and only one solution because V, by Eq. (4.4.6), is positive definite (so $\det V > 0$) if $\sigma \geq 0, \tau > 0$ and its solution $\mathbf{c}$ is a particular

solution to Eq. (4.4.3) and, in fact, it is the only stationary solution to Eq. (4.4.3).

It is even possible to find a useful expression for $\mathbf{c}$.

If in Eq. (4.4.16) we choose $\beta = \sqrt{\frac{2}{N}}$, we see that Eq. (4.4.18) says that $\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N-1)}$ are $(N-1)$ vectors with $N-1$ components forming an orthonormal basis in $\mathcal{R}^{N-1}$. Furthermore, these vectors are such that, by construction, [see, also, Eq. (4.4.7)]

$$a^{-1}\, V\boldsymbol{\eta}^{(h)} = \,\mu\,\omega_h^2\,\boldsymbol{\eta}^{(h)}. \tag{4.4.22}$$

Hence it follows

$$\boldsymbol{\gamma} = \sum_{k=1}^{N-1} \widehat{\gamma}(k)\boldsymbol{\eta}^{(k)}, \tag{4.4.23}$$

$$\mathbf{c} = \sum_{k=1}^{N-1} \widehat{c}(k)\boldsymbol{\eta}^{(k)}, \tag{4.4.24}$$

where the $\widehat{c}(k)$ are unknown and, setting $Na = L$,

$$\begin{aligned}\widehat{\gamma}(k) = (\boldsymbol{\eta}^{(k)} \cdot \boldsymbol{\gamma}) = &\sqrt{\frac{2}{N}}\Big\{\Big(\sum_{j=1}^{N-1} \mu g(ja) \sin\frac{\pi\, k}{L}\, ja\Big) \\ &+ \frac{\tau}{a^2}(h_0 \sin\frac{\pi\, k}{L} a \,+\, h_L \sin\frac{\pi k}{L}(N-1)a)\Big\}\end{aligned} \tag{4.4.25}$$

Using Eq. (4.4.22), Eq. (4.4.21) becomes

$$\widehat{c}(k) = \mu^{-1}\omega_k^{-2}\widehat{\gamma}(k) \tag{4.4.26}$$

and provides an explicit expression for the components of $\mathbf{c}$ on the "natural basis" $\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N-1)}$.

Before stating a proposition summarizing all of the above remarks, it is useful to give a very interesting definition allowing a suggestive interpretation of Eq. (4.4.21).

3 Definition. *Let $\Omega = [0, L]$, $L/a = N =$ integer. Define the "finite differences Laplace operator relative to $\mathcal{Z}_a$" as the $(N-1) \times (N+1)$ matrix D associating the vector $((D\boldsymbol{\delta})_{ja})_{j=1}^{N-1}$ with the vector $\boldsymbol{\delta} = (\delta_{ja})_{j=0}^{N}$ so that*[3]

$$(D\boldsymbol{\delta})_{ja} = \frac{\delta_{ja+a} - 2\delta_{ja} + \delta_{ja-a}}{a^2}, \qquad j = 1, \ldots, N-1. \tag{4.4.27}$$

[3] The matrix elements of D are $D_{i,j} = -\frac{2}{a^2}\delta_{i,j} + \frac{\delta_{i,j+1}+\delta_{i,j-1}}{a^2}$, $i = 1 \ldots, N-1$, $j = 0, \ldots, N$.

In this notation, Eq. (4.4.21) can be written as

$$\begin{aligned} &(\sigma\mathbf{c} - \tau D\mathbf{c})_\xi = \mu\, g(\xi), && \xi \in \Omega_a/\partial\Omega, \\ &\mathbf{c}_\xi = h_\xi, && \xi \in \partial\Omega. \end{aligned} \tag{4.4.28}$$

4 Definition. *Equation (4.4.28) for the vector* $\mathbf{c}$ *will be called, for* $\sigma \geq 0, \tau > 0$*, a "discrete non homogeneous Dirichlet problem for the region* Ω *on* $\mathcal{Z}^1_a$ *with boundary datum* h*, interior datum* μg*".*

The already remarked existence and uniqueness of the solutions of Eq. (4.4.21) can be phrased as follows.

8 Proposition. *Equation (4.4.28) admits one and only one solution for arbitrarily given boundary and interior data and for all* $\sigma \geq 0, \tau > 0$*.*

Concluding this section its results are summarized by:

9 Proposition. *The motions associated with Eqs. (4.4.1) and (4.4.2) have the form*

$$y^{(a)}_\xi(t) = c^{(a)}_\xi + \sum_{h=1}^{N-1} A_h \sqrt{\frac{2}{N}}\,\big(\sin\frac{\pi\, h}{L}\xi\big)\cos(\omega_h t + \varphi_h) \tag{4.4.29}$$

for $\xi \in \Omega_a$ *with*

$$\omega_h = \sqrt{\frac{\sigma}{\mu} + \frac{\tau}{\mu}\frac{2(1-\cos\frac{\pi\, h}{Na}a)}{a^2}}. \tag{4.4.30}$$

and the vector $\mathbf{c}^{(a)} = (c_\xi)_{\xi\in\Omega_a}$ *is the solution to the Dirichlet problem (4.4.28) with boundary datum* h *and interior datum* g*. The vector* $\mathbf{c}^{(a)}$ *is given by*

$$\begin{aligned} c^{(a)}_\xi = \sum_{k=1}^{N-1} \frac{\sin\frac{\pi}{L}k\xi}{\omega_k^2}\Big\{\Big(\frac{2}{N}\sum_{\xi'\in\Omega_a} g(\xi')\sin\frac{\pi}{L}k\xi'\Big) \\ + \frac{\tau}{a^2}\big(h_0\sin\frac{\pi\, k}{L}a \,+\, h_L\sin\frac{\pi k}{L}(N-1)a\big)\Big\} \end{aligned} \tag{4.4.31}$$

Observation. The normal modes have a remarkable "spatial structure", i.e., a remarkable ξ dependence. They are in fact interpolated by sinusoidal functions with "two nodes", i.e., two zeros, at the "extremes of the string", 0 and L, and in the h-th normal mode such a function has exactly $(h-1)$ other nodes in $[0, L]$. This is a complete description of the "wave-form" of the modes.

4.5 The Vibrating String as a Limiting Case of a Chain of Oscillators. The Case of Vanishing *g* and *h*. Wave Equation

The motivation for the choice of the Lagrangian (4.4.1) and (4.4.2) lies in the request that the mechanical system described by it be a good model for the oscillations of an elastic string.

In this section it will be shown in a mathematically precise sense how this property is actually realized in the models of Eqs. (4.4.1) and (4.4.2) when g and h vanish. We shall suppose $\sigma \geq 0, \tau > 0$.

To get an idea of what to try to prove, remark first that Eq. (4.4.3) has a formal limit given by

$$\mu \frac{\partial^2 y_\xi}{\partial t^2} = \mu g_\xi - \varrho y_\xi - \tau \frac{\partial^2 y_\xi}{\partial \xi^2} \tag{4.5.1}$$

$$y_0 = h_0, \qquad y_L = h_L, \tag{4.5.2}$$

as $a \to 0$, while Eq. (4.4.28) for the "center" of the oscillations becomes, still formally,

$$\sigma c_\xi - \frac{d^2}{d\xi^2} c_\xi = \mu g(\xi), \quad \xi \in [0, L], \qquad c_0 = h_0,\ c_L = h_L. \tag{4.5.3}$$

Hence the following proposition should look natural..

10 Proposition. *Let* $t \to \mathbf{y}^{(a)}(t)$*,* $t \in \mathcal{R}$*, be the solution of Eq. (4.4.3) with* $g = h = 0$*,* $\sigma \geq 0, \tau > 0, \mu > 0$*, following the initial datum*

$$y_{ja}^{(a)}(0) = u_0(ja), \qquad j = 1, \ldots, N-1 \tag{4.5.4}$$

$$\dot{y}_{ja}^{(a)}(0) = v_0(ja), \qquad j = 1, \ldots, N-1 \tag{4.5.5}$$

where $u_0, v_0 \in C_0^\infty((0, L)) \equiv \{$*functions in* $C^\infty([0, L])$ *vanishing in a neighborhood of* 0 *and* $L\}$*. Then,* $\forall\, t \in \mathcal{R}$*,* $\forall\, x \in [0, L]$*, the limit*

$$\lim_{\substack{a \to 0 \\ \xi \to x}} y_\xi^{(a)}(t) = w(x, t) \tag{4.5.6}$$

exists and defines a C^∞ *function on* $[0, L] \times \mathcal{R}$*, verifying the equations:*

$$\mu \frac{\partial^2 w}{\partial t^2} - \tau \frac{\partial^2 w}{\partial x^2} + \sigma w = 0, \qquad \forall\, (x, t) \in [0, L] \times \mathcal{R} \tag{4.5.7}$$

$$w(x, 0) = u_0(x), \qquad \forall\, x \in [0, L], \tag{4.5.8}$$

$$\frac{\partial w}{\partial y}(x, 0) = v_0(x), \qquad \forall\, x \in [0, L], \tag{4.5.9}$$

$$w(0, t) = 0 = w(L, t), \qquad \forall\, t \in \mathcal{R}. \tag{4.5.10}$$

Equations (4.5.7)-(4.5.10) admit one and only one C^∞ solution: this solution is explicitly given by Eq. (4.5.19) below.

Observations.
(1) This proposition makes precise the fact that a "regular" initial datum evolves through Eq. (4.4.3) into a "regular configuration". Furthermore, it explains why Eq. (4.5.7) is called the "wave equation" describing the oscillations of a string with density μ, tension τ, and restoring constant σ. In the case $\sigma = 0$, Eq. (4.5.7) is the "D'Alembert wave equation" for the vibrating string oscillating under the only action of its tension τ.
(2) The derivation of the wave equation presented here and its theory, as expressed by Proposition 10, starting from the theory of harmonic oscillators, is a celebrated theorem of Lagrange.
(3) Another explicit solution to Eqs. (4.5.7)-(4.5.10) can be found in Problem 11, p.270, (see, also, §4.7).

PROOF. Write Eq. (4.4.29) as

$$y_\xi^{(a)} = \sum_{h=1}^{N-1} \Big\{ \widetilde{A}_h \sqrt{\frac{2}{N}} \sin \frac{\pi\, h}{L} \xi \,\cdot \cos \omega_h t + \widetilde{B}_h \sqrt{\frac{2}{N}} \sin \frac{\pi\, h}{L} \xi \,\cdot \sin \omega_h t \Big\}, \quad (4.5.11)$$

where $\xi = ia$, $i = 1, \ldots, N-1$ and try to determine $\widetilde{A}_h, \widetilde{B}_h$ by imposing the initial data.

Consider the initial data of Eqs. (4.5.4) and (4.5.5) as $(N-1)$-component vectors and express them as linear combinations with suitable coefficients, of the vectors $\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(N-1)}$ with components $(\boldsymbol{\eta}^{(h)})_i = \sqrt{\frac{2}{N}} \sin \frac{\pi i a\, h}{L}$, which (as seen in §4.4) form an orthogonal basis in $\mathcal{R}^{N-1}$ [see Eqs. (4.4.16) and (4.4.18)]:

$$\begin{aligned} u_0(\xi) &= \sum_{h=1}^{N-1} \widehat{u}_0(h) \sqrt{\frac{2}{N}} \sin \frac{\pi h}{L} \xi, \qquad \xi = iq,\ i = 1, \ldots, N-1, \\ v_0(\xi) &= \sum_{h=1}^{N-1} \widehat{v}_0(h) \sqrt{\frac{2}{N}} \sin \frac{\pi h}{L} \xi, \qquad \xi = iq,\ i = 1, \ldots, N-1. \end{aligned} \quad (4.5.12)$$

After Eq. (4.5.12), it becomes immediate to impose the initial data of Eqs. (4.5.4) and (4.5.5) to Eq. (4.5.11):

$$\widetilde{A}_h = \widehat{u}_0(h), \qquad \widetilde{B}_h = \frac{\widehat{v}_0(h)}{\omega_h} \quad (4.5.13)$$

Since, on the other hand, $\widehat{u}_0(h)$ and $\widehat{v}_0(h)$ can be obtained by scalar multiplication of the vectors of Eqs. (4.5.4) and (4.5.5) by $\boldsymbol{\eta}^{(h)}$, Eq. (4.5.13) yields

$$\sqrt{\frac{2}{N}}\widetilde{A}_h = \frac{2}{N}\sum_{\xi}\big(\sin\frac{\pi h}{L}\xi\big)u_0(\xi), \tag{4.5.14}$$

$$\sqrt{\frac{2}{N}}\widetilde{B}_h = \frac{1}{\omega_h}\frac{2}{N}\sum_{\xi}\big(\sin\frac{\pi h}{L}\xi\big)v_0(\xi), \tag{4.5.15}$$

and $\sum_\xi$ runs over $\xi = ia$, $i = 1, \ldots, N-1$.

Then, by the assumptions on u_0 and v_0, Eqs. (4.5.14) and (4.5.15) contain summations over ξ which, after being multiplied by a, are the Riemann sums for the integrals between 0 and L of the functions $x \to (\sin\frac{\pi h}{L}x)u_0(x)$ and $x \to (\sin\frac{\pi h}{L}x)v_0(x)$, $x \in [0, L]$. Hence,

$$\lim_{a\to 0}\sqrt{\frac{2}{N}}\widetilde{A}_h = \frac{2}{L}\int_0^L u_0(x)\big(\sin\frac{\pi h}{L}x\big)\,dx, \tag{4.5.16}$$

$$\lim_{a\to 0}\sqrt{\frac{2}{N}}\widetilde{B}_h = \frac{1}{\overline{\omega}_h}\frac{2}{L}\int_0^L v_0(x)\big(\sin\frac{\pi h}{L}x\big)\,dx, \tag{4.5.17}$$

where, for $h = 1, 2, \ldots$ [see Eq. (4.4.30)],

$$\overline{\omega}_h = \lim_{a\to 0}\omega_h = \sqrt{\frac{\sigma}{\mu} + \frac{\tau}{\mu}(\frac{\pi h}{L})^2} \tag{4.5.18}$$

Hence, we see that the sum (4.5.11), thought of as a series in h (with vanishing terms for $h \geq N$), converges term by term, as $a \to 0$ and $\xi \to x \in [0, L]$, to the series

$$\begin{aligned} w(x,t) = \sum_{h=0}^{\infty}\sin\frac{\pi h}{L}x\Big\{&(\frac{2}{L}\int_0^L u_0(x')\sin\frac{\pi h}{L}x'\,dx')\cos\overline{\omega}(h)t \\ &+(\frac{2}{L}\int_0^L v_0(x')\sin\frac{\pi h}{L}x'\,dx')\frac{\sin\overline{\omega}(h)t}{\overline{\omega}(h)}\Big\}. \end{aligned} \tag{4.5.19}$$

We now show that the series in Eq. (4.5.19) is uniformly convergent in t and x and defines a function w verifying Eqs. (4.5.7)-(4.5.10). This will mean that a function w verifying Eqs. (4.5.7)-(4.5.10) does exist. Then we shall prove Eq. (4.5.6), and the proof will finally be concluded by proving the uniqueness of the solution to Eqs. (4.5.7)-(4.5.10).

All of the above deductions are based on the following lemma, a corollary to the Fourier theorem, proved in Appendix I, p.536.

11 Lemma. *Let $\overline{C}^\infty([0, L])$ be the set of the $C^\infty([0, L])$ real functions vanishing together with all their even derivatives in the points* 0 *and* L. *Set*

$$\overline{u}_k = \frac{2}{L}\int_0^L u(x')\,\sin\frac{\pi k}{L}x'\,dx' \tag{4.5.20}$$

$\forall\, u \in \overline{C}^{\infty}([0,L])$; then it follows that:
(i) $\forall\, \alpha > 0$, $\exists C_\alpha$ such that

$$|\overline{u}_k| \le C_\alpha\,(1+k^\alpha)^{-1}, \qquad \forall\, k = 1,2,\dots \tag{4.5.21}$$

$$(ii) u(x) = \sum_{h=0}^{\infty} \overline{u}_k \sin\frac{\pi h}{L}x \tag{4.5.22}$$

(iii) Equation (4.5.22) can be differentiated term by term an arbitrary number of times, giving rise to uniformly convergent series.
(iv) Every function of the form of Eq. (4.5.22) with $\overline{u}_k$ verifying Eq. (4.5.21) is in $\overline{C}^{\infty}([0,L])$.

Observation. Clearly $\overline{C}^{\infty}([0,L]) \supset C_0^{\infty}((0,L))$.

The proof of Proposition 10 can be continued as follows.

The uniform convergence in t and x of Eq. (4.5.19), as well as the admissibility of its term-by-term differentiations, follow from (i), Eq. (4.5.21). Call w the sum of the series (4.5.19): it verifies Eq. (4.5.7) because every term of Eq. (4.5.19) does [see Eq. (4.5.18) and do a direct check].

Equation (4.5.10) holds since $\sin\frac{\pi h}{L}x$ vanishes in 0 and in L, for all integers h. Equations (4.5.8) and (4.5.9) can be checked by computing $w(x,0)$ and $\frac{\partial w}{\partial t}(x,0)$, from Eq. (4.5.19), using (ii) of Lemma 11.

It remains to prove Eq. (4.5.6) and uniqueness. Since Eq. (4.5.11), thought of as a series in h by setting $\widetilde{A}_h, \widetilde{B}_h \equiv 0$ for $h \ge N$, converges term by term to the function in Eq. (4.5.19), we simply have to show that the series (4.5.11) is uniformly convergent in a and ξ (or, what amounts to the same, in N and ξ). It suffices to show that given $\alpha > 0$ there exists C'_α such that

$$\sqrt{\frac{2}{N}}|\widetilde{A}_h| \le \frac{C'_\alpha}{1+h^\alpha}, \qquad h = 1,2,\dots \tag{4.5.23}$$

$$\sqrt{\frac{2}{N}}|\widetilde{B}_h| \le \frac{C'_\alpha}{1+h^\alpha}, \qquad h = 1,2,\dots \tag{4.5.24}$$

Let us, for instance, prove Eq. (4.5.23). From Eqs. (4.5.14) and (4.5.22), one obtains, $\forall\, h = 1,\dots,N-1$,

$$\begin{aligned}\sqrt{\frac{2}{N}}\widetilde{A}_h =& \frac{2}{N}\sum_{\substack{\xi=ia\\ i=1,\dots,N-1}} \sin\frac{\pi h}{L}\xi\,(\sum_{k=1}^{\infty}\overline{u}_{0k}\sin\frac{\pi k}{L}\xi)\\ =& \sum_{k=1}^{\infty}\overline{u}_{0k}\,(\frac{2}{N}\sum_{\substack{\xi=ia\\ i=1,\dots,N-1}} \sin\frac{\pi h}{L}\xi\,\sin\frac{\pi k}{L}\xi)\end{aligned} \tag{4.5.25}$$

and, by Eqs. (4.4.16) and (4.4.18), it follows that for $h = 1,\dots,N-1$ and k arbitrary (even for $k > N$),

$$\frac{2}{N}\sum_{\substack{\xi=ia\\i=1,\ldots,N-1}}\sin\frac{\pi h}{L}\xi\,\sin\frac{\pi k}{L}\xi=\delta_{k,h}-\delta_{k,2N-h}+\delta_{k,h+2N}-\ldots$$
$$=\sum_{p=0}^{\infty}\delta_{k,h+2pN}-\sum_{p=1}^{\infty}\delta_{k,2pN-h}. \tag{4.5.26}$$

Hence, by Eq. (4.5.21), $\forall\,\alpha>0,\forall\,h=1,\ldots,N-1$,

$$|\sqrt{\frac{2}{N}}\widetilde{A}_h|=|\overline{u}_{0k}-\overline{u}_{0\,2N-h}+\ldots|\le\sum_{p=0}^{\infty}|\overline{u}_{0\,h+p}|\le\sum_{p=0}^{\infty}\frac{C_\alpha}{1+(h+p)^\alpha}$$
$$\le\sum_{p=0}^{\infty}\frac{C_\alpha}{\sqrt{1+h^\alpha}}\frac{1}{\sqrt{1+p^\alpha}}=\frac{C_\alpha}{\sqrt{1+h^\alpha}}(\sum_{p=0}^{\infty}\frac{1}{\sqrt{1+p^\alpha}}) \tag{4.5.27}$$

implying Eq. (4.5.23) by the arbitrariness of α and because $\widetilde{A}_h\equiv 0$ for $h\ge N$.

To show uniqueness, it is enough to show that if $w^0\in C^\infty([0,L]\times\mathcal{R})$ and verifies Eqs. (4.5.7)-(4.5.10), with $u_0=v_0=0$, then $w^0\equiv 0$.

The idea of the proof is based on energy conservation. Equations (4.5.7)-(4.5.10) should "keep memory" of the fact that they are a formal limit of Eq. (4.4.3) and it should be possible to define, for every motion w verifying them, a function which is constant as t varies and which can be obtained as the limit $a\to 0$ of the energy expression for Eq. (4.4.3). If $y_0=y_L=0$, the energy of the motions of Eq. (4.4.3) is [see Eq. (4.3.13)]

$$E^{(a)}=\frac{a\mu}{2}\sum_{\substack{\xi=ja\\j=1,\ldots,N-1}}\dot{y}_\xi^2+\frac{a\sigma}{2}\sum_{\substack{\xi=ja\\j=1,\ldots,N-1}}y_\xi^2$$
$$+\frac{a\tau}{2}\sum_{\substack{\xi=ja\\j=0,\ldots,N-1}}\frac{(y_\xi-y_{\xi+a})^2}{a^2}, \tag{4.5.28}$$

formally becoming, in the limit $a\to 0$,

$$E(w,t)=\frac{\mu}{2}\int_0^L(\frac{\partial w}{\partial t})^2dx+\frac{\sigma}{2}\int_0^L w^2dx+\frac{\tau}{2}\int_0^L(\frac{\partial w}{\partial x})^2dx. \tag{4.5.29}$$

If we show that the solutions of Eqs. (4.5.7)-(4.5.10) in $C^\infty([0,L]\times\mathcal{R})$ are such that $E(w,t)$ remains constant as t varies, uniqueness is proved. In fact, if $w(x,0)=0$ and $\frac{\partial w}{\partial t}(x,0)=0$, then $E(0)=0$, on the other hand $E(w,t)=0\Rightarrow w(t,x)=0,\forall\,x\in[0,L]$, if $\sigma\ge 0,\tau>0$. But, the difference between two solutions of Eqs. (4.5.7)-(4.5.10) is a solution with $u_0=v_0=0$ with zero energy: hence, it vanishes identically.

To show the constancy of Eq. (4.5.29) remark that

$$\frac{d}{dt}E(w,t)=\int_0^L\big(\mu\frac{\partial w}{\partial t}\frac{\partial^2 w}{\partial t^2}+\sigma w\frac{\partial w}{\partial t}+\tau\frac{\partial w}{\partial x}\frac{\partial}{\partial x}\frac{\partial w}{\partial t}\big)\,dx \tag{4.5.30}$$

Then integrate the last term in the right-hand side by parts using $\frac{\partial w}{\partial t}(0,t) = \frac{\partial w}{\partial t}(L,t)$, by Eq. (4.5.10). Collecting the integrals into a single integral and taking Eq. (4.5.7) into account, one finds

$$\frac{dE}{dt} = \int_0^L \frac{\partial w}{\partial t}\,(\mu\frac{\partial^2 w}{\partial t^2} + \sigma w - \tau\frac{\partial^2 w}{\partial x^2}) = 0 \tag{4.5.31}$$

mbe

Observations.
(1) From the proof, one can see that the condition $u_0, v_0 \in C_0^\infty((0,L))$ has only been used to apply the Lemma 11 through the observation that $C_0^\infty((0,L)) \subset \overline{C}^\infty([0,L])$.

It is then clear that Proposition 10 can be strengthened by replacing the assumption $u_0, v_0 \in C_0^\infty((0,L))$ with the assumption $u_0, v_0 \in \overline{C}^\infty([0,L])$ and by substituting Eq. (4.5.10) with

$$w(\cdot,t) \text{ and } \frac{\partial w}{\partial t}(\cdot,t) \in \overline{C}^\infty([0,L]), \forall\, t \in \mathcal{R}. \tag{4.5.32}$$

(where $\cdot$ denotes a dummy variable; in this case, $x \in [0,L]$).

In this way the existence and uniqueness theorem for the waves equations (4.5.7)-(4.5.9) and (4.5.32) with initial datum $u_0, v_0 \in \overline{C}^\infty([0,L])$ is more satisfactory because the initial regularity condition is not modified as t evolves. In fact, from the above proof it is not possible to conclude (and it is generally false) that when the initial configuration u_0, v_0 is built with elements of $C_0^\infty(0,L))$, then also the evolved configuration at time t, $w(x,t)$, $\frac{\partial w}{\partial t}(x,t)$ consists of elements in $C_0^\infty((0,L))$ (i.e., the initial regularity is generally not preserved).
(2) One may think that $u_0, v_0 \in \overline{C}^\infty([0,L])$ is still not optimal and that, perhaps, the optimal condition could be $u_0, v_0 \in C^\infty([0,L])$ plus $u_0(0) = u_0(L) = 0$, $v_0(0) = v_0(L) = 0$. By counterexamples, it can be shown that this is not the case (see exercises). To further extend the set of the initial configurations, one has to give up C^∞ smoothness.

4.5.1 Exercises

1. Consider the wave equation for $(x,t) \in \mathcal{R}^2$

$$\frac{\partial^2 w}{\partial t^2} - c^2\frac{\partial^2 w}{\partial x^2} = 0$$

Given $u, v \in C^\infty(\mathcal{R})$, show that

$$w(x,t) = \frac{u(x+ct) + u(x-ct)}{2} + \int_{x-ct}^{x+ct} v(\xi)\frac{d\xi}{2c}$$

is a C^∞ solution verifying the initial datum (u,v).

2. In the context of Problem 1, suppose that $\frac{1}{2}\int_{-\infty}^{+\infty}(v(x)^2 + c^2(\frac{\partial u(x)}{\partial x})^2)dx < +\infty$. Show that w is the only $C^\infty(\mathcal{R}^2)$ solution "with finite energy $E = \frac{1}{2}\int_{\mathcal{R}}\left(\left(\frac{\partial w}{\partial t}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2\right)dx$" and datum (u, v). (*Hint:* Repeat the energy conservation argument at the end of the proof of Proposition 10.)

3. Find the relations between u and v, in the context of Problem 1, necessary to guarantee that w is a "purely progressive" or "purely regressive" wave, i.e., $w(x,t) = a(x - ct)$ or $w(x,t) = b(x + ct)$.

4. Let $u \in C_0^\infty((0, +\infty))$ and suppose that $u(x) = 0$, unless $x \in (a, b)$, $0 < a < b < +\infty$ and $u(x) > 0$ for $x \in (a, b)$. Let $v(x) = c\frac{du}{dx}(x)$. Show that, up to a time $t_0 > 0$, the solution w of the equation $\frac{\partial^2 w}{\partial t^2} - c^2\frac{\partial^2 w}{\partial x^2} = 0$ with initial data (u, v) is such that $w(x,t) \in \overline{C}_0^\infty((0, +\infty))$ for $t < +\infty$. (*Hint:* Use Problem 3 by noting that up to $t_0 = a/c$ the solution is $w(x,t) = u(x + ct)$.)

5. Consider the wave equation on $[0, 1]$, $\frac{\partial^2 w}{\partial t^2} - c^2\frac{\partial^2 w}{\partial x^2} = 0$, with the initial data $v_0 = -c\frac{du_0}{dx}$, $u_0(x) = x^{2n}e^{(\frac{1}{2}-x)^{-2}}$ for $0 < x < \frac{1}{2}$, $u_0(x) \equiv 0$ for $|x| \geq \frac{1}{2}$. Letting $n \geq 1$, show that up to $t_0 = \frac{1}{2c}$, the function $w(x,t) = u_0(x - ct)$ if $0 \leq x - ct \leq \frac{1}{2}$ or $w(x,t) = 0$, otherwise, is a $C^{(2n)}([0, 1])$ solution following a $C^\infty([0, 1])$ datum. Infer that the conditions $u_0, v_0 \in \overline{C}^\infty([0, L])$ in Proposition 10 cannot be replaced by the more general ones of the Observation (2), p.269. (*Hint:* Show by the same energy conservation argument at the end of the proof of Proposition 10 that there is uniqueness for the $C^{(2)}$ solutions of the wave equation, etc.)

6. Is the condition $\tau > 0$ in Proposition 10 essential? If yes, give a physical interpretation of the reason.

7. A solution to the equation $\frac{\partial^2 w}{\partial t^2} - c^2\frac{\partial^2 w}{\partial x^2} + m^2 w = 0$, $(x,t) \in \mathcal{R}^2$, having the form $e^{i(kx \pm ct)}$ is called a "plane wave" solution. Its real and imaginary parts are called "real plane waves" solutions. Find the plane wave solutions to the above equation.

8. Find the energy per unit length of a real plane wave solution to the equation in Problem 7. (*Hint:* $E \overset{def}{=} \lim_{L\to\infty}\frac{1}{2L}\int_{-L}^{L}\left((\frac{\partial w}{\partial t})^2 + c^2(\frac{\partial w}{\partial x})^2 + m^2 w^2\right)dx \ldots$.)

9. Formulate and prove Proposition 10 in the case when the segment $[0, L]$ is replaced by a closed circle, i.e., the oscillators in Fig. 4.1 are ideally bound to the set of equispaced lines orthogonal to a circle with radius R, obviously without fixed extreme oscillators ("periodic boundary conditions"). Show that Eqs. (4.5.7)-(4.5.9) remain the same while Eq. (4.5.10) is replaced by $u_0, v_0 \in C^\infty(\mathcal{T}^1(2\pi R)) \overset{def}{=} C^\infty$ periodic functions with period $2\pi R$. (*Hint:* The ordinary Fourier theorem replaces Lemma 11 in the proof (which actually becomes easier).)

10. In the context of Problem 1, call $V_0(x) = \int_0^x v_0(\xi)d\xi$. Show that to compute w at the point (x,t), it is enough to know the data u_0, V_0 at the points $x \pm ct$ ("propagation along characteristic lines").

11. Consider the wave equations (4.5.7)-(4.5.10). Define $\overline{u}_0, \overline{v}_0$ as

$$\begin{aligned}
\overline{u}_0(x) &= u_0(x), \qquad \text{if} \quad 0 \leq x \leq L,\\
u_0(L + x) &= -\,u_0(L - x), \qquad \text{if} \quad L \leq L + x \leq 2L, \quad \text{and}\\
u_0(x) &= u_0(x - 2kL), \qquad \text{if} \quad x - 2kL \in [0, 2L].
\end{aligned}$$

Likewise, define $\overline{v}_0$. Show that $\overline{u}_0, \overline{v}_0$ are $C^\infty(\mathcal{R})$ functions if and only if $u_0, v_0 \in \overline{C}^\infty([0, L])$. Let $V_0(x) \overset{def}{=} \int_0^x \overline{v}_0(\xi)d\xi$. Show that the solution to Eqs. (4.5.7)-(4.5.10) can be written

$$w(x,t) = \frac{\overline{u}_0(x-ct)+\overline{u}_0(x+ct)}{2} + \frac{\overline{V}_0(x-ct)+\overline{V}_0(x+ct)}{2}$$

(see, also, §4.7). Find a statement analogous to the one in Problem 10 in terms of $\overline{u}_0, \overline{V}_0$.

4.6 Vibrating String: General Case. Dirichlet Problem in [0, L]

Having in mind the results of §4.4, it is convenient to study preliminarily what happens to the stationary solution $\mathbf{c}^{(a)}$ [see Eqs. (4.4.29) and (4.4.31)] in the limit $a \to 0$, $\xi \to x$.

The heuristic considerations at the beginning of §4.5 suggest the following proposition.

12 Proposition. *The stationary solution* $\mathbf{c}^{(a)}$ *of the oscillator-chain equations (4.4.1) and (4.4.2) given by Eq. (4.4.31) is such that the limit*

$$c(x) = \lim_{\substack{a\to 0\\ \xi\to x}} c_\xi^{(a)} \tag{4.6.1}$$

exists for $x \in [0,L]$ *and defines a function* $c \in C^\infty([0,L])$ *such that*

$$\sigma\, c - \frac{d^2c}{dx^2} = \mu g, \qquad x \in [0,L], \tag{4.6.2}$$

$$c(0) = h_0,\ \ c(L) = h_L. \tag{4.6.3}$$

Proof. Define

$$c_\xi^{(a)1} \overset{def}{=} \sum_{k=1}^{N-1} (\sin\frac{\pi k}{L}\xi)\frac{1}{\omega_k^2}(\frac{2}{N}\sum_{\xi'} g(\xi')\,\sin\frac{\pi k}{L}\xi'), \tag{4.6.4}$$

$$c_\xi^{(a)2} \overset{def}{=} \sum_{k=1}^{N-1} (\sin\frac{\pi k}{L}\xi)\frac{1}{\mu\omega_k^2}\frac{\tau}{a^2} \tag{4.6.5}$$

for $\xi = ia$, $i = 0,1,\ldots,N, N = L/a$, and by Eq. (4.4.31),

$$\mathbf{c}^{(a)} = \mathbf{c}^{(a)1} + \mathbf{c}^{(a)2} \tag{4.6.6}$$

and $\mathbf{c}^{(a)1}$ solves Eq. (4.4.28) for $h = 0$ while $\mathbf{c}^{(a)2}$ solves it for $g = 0$.

We shall separately show the existence of the limits:

$$\lim_{\substack{a\to 0\\ \xi\to x}} c^{(a)1} = c^{(1)}(x), \tag{4.6.7}$$

$$\lim_{\substack{a\to 0\\ \xi\to x}} c^{(a)2} = c^{(2)}(x), \tag{4.6.8}$$

and that, furthermore, they define two $C^\infty([0,L])$ functions verifying Eqs. (4.6.2) and (4.6.3) with $h=0$ or $g=0$, respectively.

First study Eq. (4.6.7) using Eq. (4.6.4) as a starting point. Think of Eq. (4.6.4) as a series in k with all the terms with $k \ge N$ vanishing, then such a series converges term by term, when $\xi \to x$, $a \to 0$ to the series

$$c^{(1)}(x) = \sum_{k=1}^{\infty} (\sin \frac{\pi k}{L} x) \frac{1}{\overline{\omega}(k)^2} \big(\frac{2}{L} \int_0^L g(x') \sin \frac{\pi k}{L} x' \, dx' \big), \tag{4.6.9}$$

where $\overline{\omega}(k)^2 = \frac{\sigma}{\mu} + \frac{\tau}{\mu} (\frac{\pi k}{L})^2 = \lim_{a\to 0} \omega_k^2$ is given by Eq. (4.5.18).

If $g \in \overline{C}^\infty([0,L])$, we could infer from the Lemma 11, p.266, Eq. (4.5.21), that the above series is a uniformly convergent series, term by term indefinitely differentiable. It would then be clear that $c^{(1)}$ verifies Eqs. (4.6.2) and (4.6.3) with $h=0$ since

$$\sigma c^{(1)} - \tau \frac{d^2 c^{(1)}}{dx^2} = \sum_{k=1}^{\infty} (\sin \frac{\pi k}{L} x) \cdot \mu \left(\frac{2}{L} \int_0^L g(x') \sin \frac{\pi k}{L} x' \, dx' \right) \tag{4.6.10}$$

and by Lemma 11 the right-hand side is just μg.

It would also be easy to prove the validity of Eq. (4.6.7) with $c^{(1)}$ defined by Eq. (4.6.9). One should repeat, word by word, the §4.5 proof where the convergence of $y_\xi^{(a)}(t)$ to its "term-by-term limit", Eq. (4.5.19), is discussed.

In the present case, however, $g \in C^\infty([0,L])$ but not necessarily $g \in \overline{C}^\infty([0,L])$, and the proof of Eq. (4.6.7), of the convergence of Eq. (4.6.9), and of the $C^\infty([0,L])$ nature of $c^{(1)}$ is more delicate.

Technically, such a problem must be present and it takes place because the series (4.6.10) cannot converge too well to $g(x)$: if, in fact, it did converge absolutely and if it had g as its sum, it would follow $g(0) = g(L) = 0$, for instance, which might be false for a given g. This phenomenon always appears, whenever one tries to approximate a function g with functions (in our case $\sin \frac{\pi k}{L} x$ with properties too different from those of g (for instance, $g(0) \neq 0$ in general, but all the approximating functions vanish in 0!).

The upcoming discussion is interesting because it illustrates how it is sometimes possible to bypass the obstacle just met: it is in fact a type of problem that often occurs in mathematical analysis.

We shall first show that the series in Eq. (4.6.9) converges to some function $c^{(1)}$ on $[0,L]$, continuous and once differentiable term by term. Then we shall show that Eq. (4.6.9) also verifies Eq. (4.6.7).

Finally, and this will be the most interesting part, we shall show that Eq. (4.6.9) verifies the Dirichlet problem, Eq. (4.6.2); and this will imply, by the regularity theorem, Proposition 1, p.14, that, actually, $c^{(1)} \in C^\infty([0,L])$, although, of course, it may be that $c^{(1)} \notin \overline{C}^\infty([0,L])$.

To show that the series (4.6.9) is convergent and once differentiable term by term, we can remark that, setting $g' = \frac{dg}{dx}$:

$$\begin{aligned}\overline{g}_k =&\frac{2}{L}\int_0^L g(x')\sin\frac{\pi k}{L}x'\,dx' = \Big[\frac{-1}{\pi k/L}\frac{2}{L}g'(x')\cos\frac{\pi k}{L}x'\Big]_0^L\\ &+\frac{L}{\pi k}\frac{2}{L}\int_0^L g'(x')\cos\frac{\pi k}{L}x'\,dx' = \frac{2}{\pi k}[g(0)-(-1)^k g(L)]\\ &+\frac{2}{\pi k}\int_0^L g(x')\cos\frac{\pi k}{L}x'\,dx', \qquad k=1,2,\ldots\end{aligned} \tag{4.6.11}$$

This implies, if $M_{g'} = \max_{x\in[0.L]} |g'(x)|$:

$$|\overline{g}_k| \le \frac{2}{\pi k}(|g(0)| + |g(L)| + LM_{g'}) \tag{4.6.12}$$

which means that the series (4.6.9) is uniformly convergent together with its derivative series: since $\omega(k)^2$ diverges as k^2 for $k\to\infty$, in fact, such series are respectively bounded above by the convergent series [see Eq. (4.6.12)]

$$\sum_{k=1}^{\infty}\frac{|\overline{g}_k|}{\overline{\omega}(k)^2}, \quad \text{and} \sum_{k=1}^{\infty}\frac{|\overline{g}_k|}{\overline{\omega}(k)^2}\frac{\pi k}{L} \tag{4.6.13}$$

Hence, by the series differentiation theorems, Eq. (4.6.9) converges and its derivative can be computed by series differentiation and is a continuous function (as a sum of a uniformly convergent series of continuous functions).

We now show that Eq. (4.6.9) verifies Eq. (4.6.7). Since, as already observed, the term-by-term limit of Eq. (4.6.4), thought of as a series in k, is Eq. (4.6.9), it will suffice to show that such a term-by-term limit is actually correct. In other words, it will suffice to show that Eq. (4.6.4), thought of as a series in k with all the terms with $k \ge N$ vanishing, is uniformly convergent with respect to a and ξ.

We shall show this by dominating the series (4.6.4) by the series

$$\sum_{k=1}^{\infty}\frac{1}{\omega_k^2}2M_g, \qquad \text{if } M_g = \max_{x\in[0,L]}|g(x)|, \tag{4.6.14}$$

where the terms with $k \ge N$ are thought to be zero.

Recalling the form of ω_k, see Eq. (4.4.17), and using the inequality

$$2\frac{(1-\cos\varphi)}{\varphi^2} \ge \frac{4}{\pi^2} \qquad \text{if } \varphi\in[0,\pi], \tag{4.6.15}$$

we see that if $0 \le \frac{\pi k}{L}a \le \pi$:

$$\omega_k^{-2} = [\frac{\sigma}{\mu} + \frac{\tau}{\mu}\frac{2(1-\cos\frac{\pi k}{L}a)}{a^2}]^{-1} \le (\frac{\sigma}{\mu} + \frac{\tau}{\mu}\frac{4}{\pi^2 L^2}k^2)^{-1}. \tag{4.6.16}$$

Hence, Eq. (4.6.14) is a series which is bounded above by the series in which ω_k^{-2} is replaced by the right-hand side of Eq. (4.6.16), and this last series is dominated by

$$\sum_{k=1}^{\infty} 2M_g\,(\frac{\sigma}{\mu} + \frac{\tau}{\mu}\frac{4}{\pi^2 L^2}k^2)^{-1} < +\infty \tag{4.6.17}$$

having removed only in this last step the restriction $k \leq N-1$. This proves that Eq. (4.6.4) is uniformly convergent with respect to the parameters a, ξ, N and, hence, Eq. (4.6.7) follows.

We now must show that Eq. (4.6.9) is a $C^\infty([0,L])$ function verifying Eqs. (4.6.2) and (4.6.3) with $h_0 = 0, h_L = 0$. Equation (4.6.3) is obvious since Eq. (4.6.9) has been proved to converge (and all its terms vanish for $x = 0$ or $x = L$). To prove Eq. (4.6.2), we use the fact that, as already remarked, it would be obvious if $g \in \overline{C}^\infty([0,L])$.

Given $\varepsilon > 0$, let $g_\varepsilon \in C_0^\infty((0,L)) \subset \overline{C}^\infty([0,L])$ be a function such that:

$$(i) \qquad g_\varepsilon(x) = g(x) if \varepsilon \leq x \leq L - \varepsilon. \tag{4.6.18}$$

$$(ii) \qquad \frac{1}{L}\int_0^L |g_\varepsilon(x) - g(x)|dx < \varepsilon. \tag{4.6.19}$$

(iii) The derivative g'_ε of g_ε, see Fig. 4.6, is such that

$$\int_0^L |g'_\varepsilon(x)|dx \leq \int_0^L |g'(x)|dx + 2M_g. \tag{4.6.20}$$

We leave as an exercise based on Appendix C, p. 521, the proof that such a function indeed exists (note that (iii) expresses that g_ε can be chosen to go from zero to $g(\varepsilon)$ or from $g(L-\varepsilon)$ to 0 without oscillating too much, i.e., with a derivative changing sign once at most without growing too large).

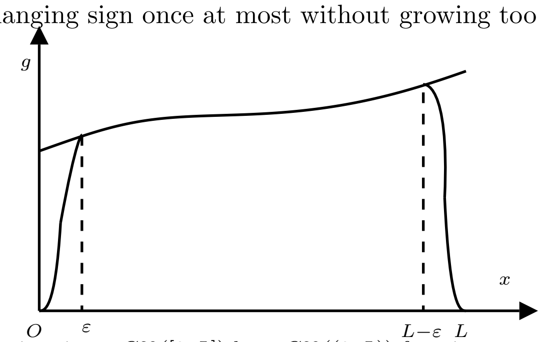


Figure 4.6: Approximating a $C^\infty([0,L])$ by a $C_0^\infty((0,L))$ function.

Then define

$$\overline{g}_{\varepsilon,k} = \frac{2}{L}\int_0^L g_\varepsilon(x')\sin\frac{\pi k}{L}x'\,dx', \quad \overline{c}_k^{(1)} = \sum_{k=1}^{\infty}\frac{\overline{g}_{\varepsilon,k}}{\overline{\omega}(k)^2}\,\sin\frac{\pi k}{L}x, \tag{4.6.21}$$

and, since $g_\varepsilon \in \overline{C}^\infty([0,L])$, we already mentioned that

$$\sigma\,\overline{c}^{(1)\varepsilon} - \tau\frac{d^2\overline{c}^{(1)\varepsilon}}{dx^2} = \mu\, g_\varepsilon, \qquad \overline{c}^{(1)\varepsilon}(0) = \overline{c}^{(1)\varepsilon}(L) = 0 \tag{4.6.22}$$

which implies

$$\begin{aligned}\overline{c}^{(1)\varepsilon}(x) \equiv & \int_0^x \frac{d\overline{c}^{(1)\varepsilon}}{dx}(x')dx' \equiv \int_0^x dx'[\frac{d\overline{c}^{(1)\varepsilon}}{dx}(0) + \int_0^{x'} \frac{d^2\overline{c}^{(1)\varepsilon}}{dx^2}(x'')dx''] \\ =& x\frac{d\overline{c}^{(1)\varepsilon}}{dx}(0) + \int_0^x dx' \int_0^{x'} dx''[\frac{\sigma\overline{c}^{(1)\varepsilon}(x'') - \mu g_\varepsilon(x'')}{\tau}].\end{aligned} \tag{4.6.23}$$

If we show that uniformly in $x \in [0,L]$:

$$c^{(1)}(x) = \lim_{\varepsilon\to 0} \overline{c}^{(1)\varepsilon}(x), \qquad \frac{dc^{(1)}}{dx}(x) = \lim_{\varepsilon\to 0}\frac{d\overline{c}^{(1)\varepsilon}}{dx}(x) \tag{4.6.24}$$

we shall be able to take the limit in Eq. (4.6.23) and obtain

$$c^{(1)}(x) = \frac{dc^{(1)}}{dx}(0) + \int_0^x dx' \int_0^{x'} dx''[\frac{\sigma c^{(1)}(x'') - \mu g_\varepsilon(x'')}{\tau}], \tag{4.6.25}$$

implying by assumed continuity of g and by the above proved continuity of $c^{(1)}$ that $c^{(1)}$ is twice differentiable and by twofold differentiation of Eq. (4.6.25) that it verifies Eq.(4.6.2).

The regularity theorem of §2.2, Proposition 1, p.14, will then permit us to deduce from the fact that $c^{(1)}$ is twice differentiable with continuous derivatives and verifies Eq. (4.6.2) that $c^{(1)}$ is in $C^\infty([0,L])$.[4]

Therefore, it remains to prove that the limits of Eq. (4.6.24) are correct and uniform in $x \in [0,L]$.

We already know that $c^{(1)}$ and its first derivative are given by the series (4.6.9) and by the sum of its term-by-term derivative. Such series are also the limits, term-by-term, of the series in Eq. (4.6.21) and of its derivative series because by Eqs. (4.6.19) and (4.6.21):

$$|\overline{g}_{\varepsilon,k} - \overline{g}_k| < 2\varepsilon, \qquad \forall\, k > 0 \tag{4.6.26}$$

Hence, the proof of Eq. (4.6.24) is again a problem of exchanging a limit with a series summation.

The necessary uniformity of the limit and the convergence of the series follow from the identity:

[4] This also follows directly from Eq. (4.6.2) since it shows that the second derivative of $c^{(1)}$ is continuously differentiable because such are g and $c^{(1)}$, etc.

$$\begin{aligned}\overline{g}_{\varepsilon,k} =& \frac{2}{L}\int_0^L g_\varepsilon(x)\,\sin\frac{\pi k}{L}x\,dx = \frac{2}{\pi k}\int_0^L g'_\varepsilon(x)\cos\frac{\pi k}{L}x\,dx \\ =& \frac{2}{\pi k}\Big(\int_\varepsilon^{L-\varepsilon} g'(x)\cos\frac{\pi k}{L}x\,dx + \int_{x\notin[\varepsilon,L-\varepsilon]} g'_\varepsilon(x)\cos\frac{\pi k}{L}x\,dx\Big).\end{aligned} \tag{4.6.27}$$

Hence, by Eq. (4.6.20),

$$\begin{aligned}|\overline{g}_{\varepsilon,k}| \leq& \frac{2LM_{g'} + 2\int_0^L |g'_\varepsilon(x)|dx|dx}{\pi k} \\ \leq& \frac{4LM_{g'} + 4M_g}{\pi k}\end{aligned} \tag{4.6.28}$$

and, therefore, the series (4.6.21) and its derivative series are dominated by the series (ε independent and convergent):

$$\sum_{k=1}^\infty \frac{4LM_{g'} + M_g}{\pi k \overline{\omega}(k)^2} \qquad \text{and} \sum_{k=1}^\infty \frac{\pi}{L}\frac{4LM_{g'} + M_g}{\pi k \overline{\omega}(k)^2} \qquad \text{and} \tag{4.6.29}$$

proving their uniform convergence and, hence, Eq. (4.6.24).

To conclude the proof of Proposition 12, we still have to treat $c^{(a)2}$ defined by Eq. (4.6.5) or by being the unique solution to the equations [see Eq. (4.4.28)]:

$$\begin{aligned}&(\sigma \mathbf{c}^{(a)2} - \tau D\mathbf{c}^{(a)2})_\xi = 0, \qquad \xi = ja,\ j = 1, \ldots, N-1, \\ &c_0^{(a)2} = h_0, \qquad c_L^{(a)2} = h_L.\end{aligned} \tag{4.6.30}$$

Suppose, first, that $\sigma > 0$. The expression (4.6.5) is not too helpful for investigating the limit $a \to 0, \xi \to x$. We therefore look for an alternative representation for $c^{(a)2}$ in analogy with the theory of linear differential equations.

We look for a solution of Eq. (4.6.30) having the form

$$c_{ja}^{(a)} = \beta_0 e^{-\lambda\, ja} + \beta_1 e^{-\lambda\,(L-ja)}, \qquad j = 0, \ldots, N \tag{4.6.31}$$

where in the second term we use (instead of an arbitrary constant factor β) the constant factor $\beta_1 e^{-\lambda L}$, still arbitrary because such is β_1 but yielding a more symmetric expression (in which 0 and L "play the same role").

The parameters β_0, β_1, are to be determined so that Eq. (4.6.30) is verified.

Equation (4.6.30) will hold for $j = 2, \ldots, N-2$ if

$$\sigma + \frac{2\tau}{a^2}(1 - \frac{e^{\lambda a} + e^{-\lambda a}}{2}) = 0 \tag{4.6.32}$$

which, via a simple discussion, is shown to admit a unique positive solution λ such that

$$\lim_{a\to 0} \lambda = \sqrt{\frac{\sigma}{\tau}} \equiv \lambda_0 \tag{4.6.33}$$

Furthermore, Eq. (4.6.30) for $j = 1$ or $N-1$ says, by taking Eq. (4.6.32) into account,

$$\begin{aligned} \beta_0 + \beta_1 e^{-\lambda L} &= h_0 \\ \beta_0 e^{-\lambda L} + \beta_1 &= h_L \end{aligned} \quad \Rightarrow \quad \begin{aligned} \beta_0 &= \frac{h_0 - h_L e^{-\lambda L}}{1 - e^{-2\lambda L}} \\ \beta_1 &= \frac{h_L - h_0 e^{-\lambda L}}{1 - e^{-2\lambda L}} \end{aligned} \tag{4.6.34}$$

From Eqs. (4.6.31), (4.6.33), and (4.6.34), it is now immediate to take the limit $a \to 0$, $ja \to x$. One finds

$$c^{(2)}(x) = \lim_{\substack{a \to 0 \\ ja \to x}} c_{ja}^{(a)2} = \frac{h_0 - h_L e^{-\lambda_0 L}}{1 - e^{-2\lambda_0 L}} e^{-\lambda_0 x} + \frac{h_L - h_0 e^{-\lambda L}}{1 - e^{-2\lambda L}} e^{\lambda_0 (L-x)} \tag{4.6.35}$$

which is immediately checked to verify Eqs. (4.6.2) and (4.6.3) with $g = 0$. The case $\sigma = 0$ is analogously treated by replacing Eq. (4.6.31) with

$$c_{ja}^{(a)2} = \beta_0 + \beta_1\, ja, \tag{4.6.36}$$

and one eventually finds

$$c^{(2)}(x) = h_0 + \frac{x}{L}(h_L - h_0) \tag{4.6.37}$$

and Proposition 12 is completely proved. mbe

It is useful to collect all the results of this and the preceding section into single statement.

13 Corollary. *Let $t \to \mathbf{y}^{(a)}(t)$ be a motion verifying Eq. (4.4.3) with initial data*

$$y_\xi^{(a)}(0) = c_\xi^{(a)}(0) + u_0(\xi), \qquad \dot{y}_\xi^{(a)}(0) = v_0(\xi), \tag{4.6.38}$$

where $u_0, v_0 \in \overline{C}^\infty([0, L])$ and $\mathbf{c}^{(a)}$ is a solution to the discrete Dirichlet problem, Eq. (4.4.28). Then the limit

$$c^{(2)}(x) = \lim_{\substack{a \to 0 \\ \xi \to x}} y_\xi^{(a)}(t) = c(c) + \overline{w}(x, t) \tag{4.6.39}$$

exists and $c \in C^\infty([0, L])$ is the solution to the "Dirichlet problem"

$$\sigma\, c - \tau\, \frac{d^2 c}{dx^2} = \mu g, \qquad c(0) = h_0,\ c(L) = h_L, \tag{4.6.40}$$

while $\overline{w} \in C^\infty([0, L] \times \mathcal{R})$ verifies the wave equations (4.5.7)-(4.5.10) and $\overline{w}(\cdot, t) \in \overline{C}^\infty([O, L])$, $\forall\, t \in \mathcal{R}$.

4.7 Elastic Film. The Dirichlet Problem in $\Omega \subset \mathcal{R}^2$ and General Considerations on the Waves

The theory of the oscillations of an elastic film is considerably more complex and interesting than that of the elastic string of §4.3-4.6. The results, however, are very similar. We shall not enter into the details of a theory that would lead us quite far from our program of analysis of the simplest mechanical systems.

We only give some terminology and formulate for illustrative purposes some easy propositions.

We shall then conclude our introduction to wave theory by defining the wave propagation velocity, studying it in the simple case of the elastic string subject only to tension forces ($\sigma = 0, h = 0, g = 0$).

5 Definition. *Let $\Omega \subset \mathcal{R}^2$ be a bounded open connected region with a boundary $\partial\Omega$ which is a regular surface (see Definition 10, p.170). Let $\Omega_a = \Omega \cap \mathcal{Z}_a^2$, and $\partial\Omega_a = \{$ set of points of $\partial\Omega$ lying on the intersections between $\partial\Omega$ and the bonds of the lattice $\mathcal{Z}_a\}$.*
The discrete Laplace operator on Ω relative to $\mathcal{Z}_a^2$ is defined as the linear transformation D associating with every vector $\boldsymbol{\delta} = (\delta_\xi)_{\xi \in \Omega_a \cup \partial\Omega_a}$ the vector $((D\boldsymbol{\delta})_\xi)_{\xi \in \Omega_a}$ given by

$$(D\boldsymbol{\delta})_\xi = -\sum_{\mathbf{e}} \frac{a}{\varepsilon_a(\boldsymbol{\xi}, \mathbf{e})} \frac{\delta_\xi - \delta_{\xi + \varepsilon_a(\boldsymbol{\xi},\mathbf{e})\mathbf{e}}}{a^2}, \qquad \xi \in \Omega_a, \tag{4.7.1}$$

where $\mathbf{e} = \pm\mathbf{e}_1, \pm\mathbf{e}_2$ ($\mathbf{e}_1$, and $\mathbf{e}_2$ being the two unit vectors parallel to the axes of $\mathcal{Z}_a^2$) and, for $\boldsymbol{\xi} \in \Omega_a$:

$$\begin{aligned}\varepsilon_a(\boldsymbol{\xi}, \mathbf{e}) =& \{\textit{distance between } \boldsymbol{\xi} \textit{ and its nearest neighbor in} \\ & \Omega_a \cup \partial\Omega_a \textit{ in the direction } \mathbf{e}\}\end{aligned} \tag{4.7.2}$$

The "$\mathcal{Z}_a^2$-discretized" Dirichlet problem in Ω with interior data $\mathbf{g} = (g_{\boldsymbol{\xi}})_{\boldsymbol{\xi} \in \Omega_a}$ and boundary data $\mathbf{h} = (h_{\boldsymbol{\xi}})_{\boldsymbol{\xi} \in \partial\Omega_a}$ are the equations

$$\sigma\,\delta_\xi - \tau\,(D\boldsymbol{\delta})_{\boldsymbol{\xi}} = g_{\boldsymbol{\xi}}, \quad \boldsymbol{\xi} \in \Omega_a, \tag{4.7.3}$$

$$\delta_{\boldsymbol{\xi}} = h_{\boldsymbol{\xi}}, \qquad \boldsymbol{\xi} \in \partial\Omega_a. \tag{4.7.4}$$

Using the invertibility of positive-definite matrices, Appendix F, p.525, the following proposition is checked along the same pattern of the proof of Proposition 8,§4.4, p.263.

14 Proposition. *If $\sigma \geq 0, \tau > 0$, the Dirichlet problem [Eqs. (4.7.3) and (4.7.4)] always admits one and only one solution for any given boundary and interior data.*

Again, in the same way as in §4.4 and §4.5, one may check the following proposition.

15 Proposition. *Given* $g \in C^\infty(\mathcal{R}^2)$, $h \in C^\infty(\partial\Omega)$,[5] *consider the mechanical system with Lagrangian function [see Eqs. (4.3.13) and (4.3.5)]*

$$\begin{aligned}\mathcal{L} = \sum_{\xi\in\Omega_a} \Big(\frac{\mu}{2}a^2\dot{y}_{\boldsymbol{\xi}}^2 + \mu a^2 g(\boldsymbol{\xi})y_{\boldsymbol{\xi}} - \frac{\sigma}{2}a^2 y_{\boldsymbol{\xi}}^2\Big)\\ -\frac{\tau}{2}a^2 \sum_{\xi\in\Omega_a}\sum_{\mathbf{e}} \frac{a}{\varepsilon_a(\boldsymbol{\xi},\mathbf{e})}\frac{1}{\nu(\mathbf{e},\boldsymbol{\xi})}\frac{(y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}+\varepsilon_a(\boldsymbol{\xi},\mathbf{e})\mathbf{e}})}{a^2},\end{aligned} \tag{4.7.5}$$

$$y_{\boldsymbol{\xi}} = h(\boldsymbol{\xi}), \qquad \forall\, \boldsymbol{\xi} \in \partial\Omega \tag{4.7.6}$$

This mechanical system has one and only one equilibrium configuration $\mathbf{y} = \mathbf{c}^{(a)}$. *It is described by the solution* $\mathbf{c}^{(a)}$ *of the* $\mathcal{Z}_a^2$*-discretized Dirichlet problem with interior data* $(\mu g(\boldsymbol{\xi}))_{\boldsymbol{\xi}\in\Omega_a}$ *and boundary data* $(h(\boldsymbol{\xi}))_{\boldsymbol{\xi}\in\partial\Omega_a}$.

Observation. More generally, if one is not interested in the limit $a \to 0$ the conditions , $g \in C^\infty(\mathcal{R}^2)$, $h \in C^\infty(\partial\Omega)$ can be replaced by $g = g(\boldsymbol{\xi})_{\boldsymbol{\xi}\in\Omega_a}$, and $h = (h_{\boldsymbol{\xi}})_{\boldsymbol{\xi}\in\partial\Omega_a}$.
Difficulties arise when one wishes to study the $a \to 0$ limit. Basically, one can say that the difficulties are due to the impossibility of providing the eigenvalues $\omega_1^2, \omega_2^2, \ldots$ and the respective eigenvectors $\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \ldots$, describing the normal modes of the system of Eqs. (4.7.5) and (4.7.6), in a very explicit way, as in the case $d = 1$. Hence, the theory has to be developed in a somewhat more abstract way.

An example of a result that should be possible to obtain is as follows.

16 Proposition. *The stationary solution* $\mathbf{c}^{(a)}$ *of the equations for the mechanical system of Eqs. (4.7.5) and (4.7.6) with* $g \in C^\infty(\mathcal{R}^2)$, $h \in C^\infty(\partial\Omega)$ *is such that the limit*

$$\lim_{\substack{\boldsymbol{\xi}\to\mathbf{x}\\ a\to 0}} c_{\boldsymbol{\xi}}^{(a)} = \mathbf{c}(\mathbf{x}), x \in \overline{\Omega}, \tag{4.7.7}$$

exists and defines a function $c \in C^\infty(\overline{\partial}\Omega)$ *such that*

$$\sigma\, c(\mathbf{x}) - \tau\, \Delta\, c(\mathbf{x}) = \mu\, g(\mathbf{x}), \qquad \mathbf{x} \in \Omega, \tag{4.7.8}$$

$$c(\mathbf{x}) = h(\mathbf{x}), \qquad \mathbf{x} \in \partial\Omega, \tag{4.7.9}$$

where $\Delta\, f(\mathbf{x}) = \sum_{i=1}^2 \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x})$, $\forall\, f \in C^\infty(\overline{\Omega})$. *Furthermore, Eqs. (4.7.8) and (4.7.9) have a unique solution in* $C^\infty(\overline{\Omega})$.
The motions $t \to \mathbf{y}^{(a)}(t)$, $t \in \mathcal{R}$, *of the above mechanical system, fulfilling the initial conditions*

[5] See footnote [1].

$$y_{\boldsymbol{\xi}}^{(a)}(0) = c_{\boldsymbol{\xi}}^{(a)} + u_0(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in \Omega_a \tag{4.7.10}$$

$$\dot{y}_{\boldsymbol{\xi}}^{(a)}(0) = v_0(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in \Omega_a, \tag{4.7.11}$$

with $u_0, v_0 \in C_0^\infty(\Omega)$, are such that the limit

$$\lim_{\substack{a \to 0 \\ \boldsymbol{\xi} \to \mathbf{x}}} y_{\boldsymbol{\xi}}^{(a)}(t) = w(\mathbf{x}, t), \qquad \mathbf{x}, t \in \overline{\Omega} \times \mathcal{R} \tag{4.7.12}$$

exists and defines a $C^\infty(\overline{\Omega} \times \mathcal{R})$ function. Furthermore, setting

$$w(\mathbf{x}, t) = c(\mathbf{x}) + w(\mathbf{x}, t), \tag{4.7.13}$$

it is

$$\mu\, \overline{w}(\mathbf{x}, t) - \tau\, \Delta\overline{w}(\mathbf{x}, t) + \mu \frac{\partial^2 \overline{w}}{\partial t^2}(\mathbf{x}, t) = 0, \tag{4.7.14}$$

$$\overline{w}(\mathbf{x}, 0) = u_0(\mathbf{x}), \qquad \frac{\partial^2 \overline{w}}{\partial t^2}(\mathbf{x}, 0) = v_0(\mathbf{x}) \tag{4.7.15}$$

$$\overline{w}(\mathbf{x}, 0) = 0, \ \frac{\partial \overline{w}}{\partial t}(\mathbf{x}, t) = 0, \qquad \forall\, \mathbf{x} \in \partial\Omega, \forall\, t \in \mathcal{R} \tag{4.7.16}$$

Finally, there is a family of functions $S^{(h)} \in C^\infty(\overline{\Omega})$, $h = 1, 2, \ldots$, vanishing on $\partial\overline{\Omega}$ and a sequence $\overline{\omega}(h)$, $h = 1, 2, \ldots$, of positive numbers such that

$$\overline{w}(\mathbf{x}, t) = \sum_{h=1}^{\infty} S^{(h)}(\mathbf{x}) \left(\widehat{u}(h) \cos \overline{\omega}(h) t + \frac{\widehat{v}(h)}{\overline{\omega}(h)} \sin \overline{\omega}(h) t\right), \tag{4.7.17}$$

where

$$\widehat{u}(h) = \int_\Omega S^{(h)}(\mathbf{x})\, u_0(\mathbf{x}) d\mathbf{x}, \qquad \widehat{v}(h) = \int_\Omega S^{(h)}(\mathbf{x})\, v_0(\mathbf{x}) d\mathbf{x}, \tag{4.7.18}$$

and the series Eq. (4.7.17) converges, $\forall\, \mathbf{x} \in \overline{\Omega}$, $\forall\, t \in \mathcal{R}$.

Observations.
(1) The analogy between the vibrating string and the vibrating film would then be essentially complete. However, this author does not know if there is a proof of Proposition 16 (admitting its truth) in the above generality.
(2) There is a case in which an obvious variation of the above proposition holds and its proof is very simple. It is the case in which Ω is a torus (i.e., Ω is a "bicycle tire") and $\sigma > 0$. Mathematically, this is the system associated with the Lagrangian that follows; let $N = L/a =$ integer, $Q_L = [0, L-a] \times [0, L-a]$:

$$\begin{aligned}\mathcal{L}_{per} =&\mu a^2 \sum_{\boldsymbol{\xi}\in Q_L\cap\mathcal{Z}_a^2} \big(\frac{1}{2}\dot{y}_{\boldsymbol{\xi}}^2 + g(\boldsymbol{\xi})y_{\boldsymbol{\xi}}\big)\\ &-\frac{\sigma}{2}a^2 \sum_{\boldsymbol{\xi}\in Q_L\cap\mathcal{Z}_a^2} y_{\boldsymbol{\xi}}^2 - \frac{\tau}{2}a^2 \sum_{\boldsymbol{\xi}\in Q_L\cap\mathcal{Z}_a^2}\sum_{\mathbf{e}} \frac{(y_{\boldsymbol{\xi}} - y_{\boldsymbol{\xi}+a\,\mathbf{e}})^2}{a^2}\end{aligned} \tag{4.7.19}$$

and in the last sum the points which do not belong to $Q_L \cap \mathcal{Z}_a^2$ and which correspond to the points adjacent to the boundary ∂Q_L have to be identified with the points on ∂Q_L opposite to them.

In other words, $Q_L \cap \mathcal{Z}_a^2$ is thought of as a "discrete torus" and the film looses its boundary, becoming a "tube".

The theory of Eq. (4.7.19) is identical to that of the vibrating string. Actually it is technically even easier (and analogous to Problem 9, §4.5, on the vibrating string). The role played by the functions $\sqrt{\frac{2}{N}}\,\sin(\frac{\pi k}{L}ja)$ in the vibrating-string case is now played by

$$S^{(h_1,h_2)}(\boldsymbol{\xi}) = \frac{1}{N}e^{\frac{2\pi i}{L}(j_1ah_1+j_2ah_2)} \qquad \text{if } \boldsymbol{\xi} = (j_1a, j_2a). \tag{4.7.20}$$

with $(j_1, j_2) \in \mathcal{Z}^2$, integers. The ω_h^2 is now replaced by

$$\omega_{h_1h_2}^2 = \frac{\sigma}{\mu} + \frac{\tau}{\mu}2[\frac{1-\cos\frac{2\pi h_1}{L}a}{a^2} + \frac{1-\cos\frac{2\pi h_2}{L}a}{a^2}], \tag{4.7.21}$$

while the role of Lemma 11, §4.5, is simply played by the two-dimensional Fourier theorem.

The detailed development of the theory of the motion of Eq. (4.7.19) (and of the analogous one-dimensional system, Problem 11, §4.5) is a very useful exercise. The reader will however realize that the assumption $\sigma > 0$ cannot, in the case of such periodic boundary conditions, be replaced by $\sigma \geq 0$ (which is the physical meaning of this?)

To conclude our analysis of the ordered systems of oscillators, we define and study concisely the notion of velocity of wave propagation.

6 Definition. *Let Ω be an open region with regular boundary $|dpr\Omega$, $\Omega \subset \mathcal{R}^d$, $d = 1$ or $d = 2$.*
Consider the wave equation in Ω for $w \in C^\infty(\Omega \times \mathcal{R})$:

$$\mu\frac{\partial^2 w}{\partial t^2} - \tau\,\Delta w + \sigma w = 0. \qquad (\mathbf{x}, t) \in \Omega \times \mathcal{R}, \tag{4.7.22}$$

$$w(\mathbf{x}, t) = 0 = \frac{\partial w}{\partial t}(\mathbf{x}, t), \qquad \mathbf{x}x \in \Omega \tag{4.7.23}$$

with initial data

$$w(\mathbf{x}, 0) = u_0(\mathbf{x}), \tag{4.7.24}$$

$$\frac{\partial w}{\partial t}(\mathbf{x}, 0) = v_0(\mathbf{x}) \tag{4.7.25}$$

and suppose $u_0, v_0 \in C_0^\infty(\Omega)$ and vanishing outside a neighborhood with radius ε around $\mathbf{x}_0 \in \Omega$. Given $\mathbf{x}_1 \in \Omega, \mathbf{x}_1 \neq \mathbf{x}_0$, let

$$\begin{aligned} t_\varepsilon(\mathbf{x}_0, \mathbf{x}_1) = \inf_{u_0, v_0} \{\inf \ &\textit{of the set of the values } t \textit{ for which there} \\ &\textit{is } t' < t,\ t' > 0, \textit{ when } w(\mathbf{x}_1, t') \neq 0. \end{aligned} \tag{4.7.26}$$

Obviously, $t_\varepsilon(\mathbf{x}_0, \mathbf{x}_1) > t_{\varepsilon'}(\mathbf{x}_0, \mathbf{x}_1)$ if $\varepsilon' > \varepsilon$, and

$$t(\mathbf{x}_0, \mathbf{x}_1) = \sup_{\varepsilon > 0} t_\varepsilon(\mathbf{x}_0, \mathbf{x}_1) \tag{4.7.27}$$

is the "minimum time" needed for a perturbation of the equilibrium, (i.e., flat), string, or film, initially located around $\mathbf{x}_0$ to "reach" $\mathbf{x}_1$.
The "wave velocity" of the waves described by Eqs. (4.7.22) and (4.7.23) is naturally defined as

$$C = \sup_{\mathbf{x}_1 \neq \mathbf{x}_0} \frac{|\mathbf{x}_1 - \mathbf{x}_0|}{t(\mathbf{x}_0, \mathbf{x}_1)}. \tag{4.7.28}$$

Observation. In the $d = 2$ case, we did not prove existence and uniqueness theorems for Eqs. (4.7.22)-(4.7.25), while for $d = 1$ we did. However, if we set $t_\varepsilon(\mathbf{x}_0, \mathbf{x}_1) = +\infty$ if for every (u_0, v_0) there is no solution to Eqs. (4.7.22)-(4.7.25) and if, in case of non unique solutions, we take into account all the solutions in the infimum in Eq. (4.7.26), the above definition also makes sense for $d = 2$.
In any case, this is not a real problem since existence and uniqueness for Eqs. (4.7.22)-(4.7.25) for $u_0, v_0 \in C^\infty(\Omega)$ can be proved in a satisfactory sense.

Let us prove the following proposition for $\sigma = 0$:

17 Proposition. *Let $d = 1$, $\Omega = (0, L)$. The wave propagation velocity of the waves described by Eqs. (4.7.22) and (4.7.23) with $\sigma \geq 0, \tau > 0, \mu > 0$ is*

$$C = \sqrt{\frac{\tau}{\mu}} \tag{4.7.29}$$

independent on the value of σ.

Proof. (Case $\sigma = 0$ only). From Eq. (4.5.19), we derive by trigonometry:

$$
\begin{aligned}
w(x,t) &= \sum_{h=1}^{\infty} \sin\frac{\pi h}{L}x \left[\widehat{u}_0(h)\,\cos\frac{\pi h}{L}Ct \,+\, \frac{\widehat{v}_0(h)}{C\frac{\pi h}{L}}\sin\frac{\pi h}{L}Ct\right] \\
&= \sum_{h=1}^{\infty} \frac{\widehat{u}_0(h)}{2}\big[\sin\frac{\pi h}{L}(x+Ct)+\sin\frac{\pi h}{L}(x-Ct)\big] \\
&\quad + \sum_{h=1}^{\infty} \frac{\widehat{v}_0(h)}{2C\frac{\pi h}{L}}\big[\cos\frac{\pi h}{L}(x+Ct)+\cos\frac{\pi h}{L}(x-Ct)\big],
\end{aligned}
\tag{4.7.30}
$$

since $\overline{\omega}_h^2 = C^2(\frac{\pi h}{L})^2$. Then, let $\forall\, x \in \mathcal{R}$,

$$
u_0^*(x) = \sum_{h=1}^{\infty} \widehat{u}_0(h)\sin\frac{\pi h}{L}x, \qquad v_0^*(x) = \sum_{h=1}^{\infty} \widehat{v}_0(h)\sin\frac{\pi h}{L}x, \tag{4.7.31}
$$

and, by the Lemma 11, p.266, plus the periodicity and parity properties of the sine:

$$
\begin{aligned}
&(i)\quad u_0^*(x) = u_0(x),\ v_0^*(x) = v_0(x), \qquad \forall\, x \in [0,L],\\
&(ii)\ \ u_0^*(L+x) = -u_0(L-x),\ \ v_0^*(L+x) = -v_0(L-x),\ \ \forall\, x \in [0,L],
\end{aligned}
\tag{4.7.32}
$$

(iii) u_0^*, v_0^* are periodic C^∞ functions with period $2L$, i.e., u_0^*, v_0^* are obtained from u_0, v_0, by first reflecting them about L and then by periodic continuation of the function on $[0, 2L]$ thus constructed.

If u_0 has support in a neighborhood with radius εaround x_0, one finds that u_0 is described in Fig. 4.7.

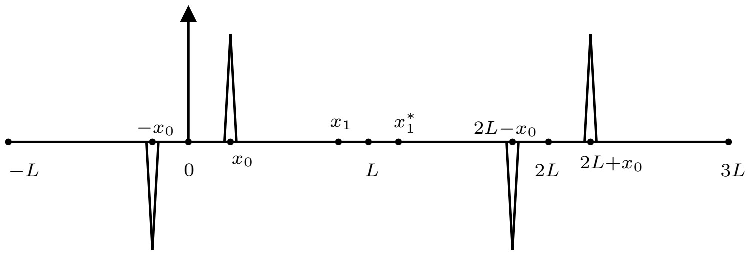


Figure 4.7.Example of graph of u_0^*.

Equation (4.7.30) can be written in terms of u_0^*, v_0^* as

$$
w(x,r) = \frac{1}{2}[u_0^*(x+Ct)+u_0^*(x-Ct)] - \frac{1}{2C}\int_{x-Ct}^{x+Ct} v_0^*(\xi)d\xi \tag{4.7.33}
$$

[see, also, Problem 11, §4.5, for an alternative proof of Eq. (4.7.33)].

Then, for instance, one sees from the picture that in order that $u_0^*(x{-}Ct) \neq 0$ the point $x_1 - Ct$ has to fall inside some of the intervals where u_0^* is not zero; hence, t must be such that

$$\frac{|x_1 - x_0| + \varepsilon}{C} \geq t \geq \frac{|x_1 - x_0| - \varepsilon}{C} \tag{4.7.34}$$

and a similar bound on t can be found likewise, discussing the the third terms in Eq. (4.7.33). This clearly implies Eq. (4.7.29). mbe

4.8 Anharmonic Oscillators. Small Oscillations and Integrable Systems

Consider an ℓ-degrees-of-freedom system with Lagrangian function

$$\mathcal{L}(\dot{\boldsymbol{\beta}}, \boldsymbol{\beta}) = \sum_{i,j=1}^{\ell} \frac{1}{2} g_{i,j}(\boldsymbol{\beta}) \dot{\beta}_i \dot{\beta}_j \;-\; V(\boldsymbol{\beta}) \tag{4.8.1}$$

where g is a given $C^\infty(\mathcal{R}^\ell)$ positive-definite $\ell \times \ell$ matrix and $V \in C^\infty(\mathcal{R}^\ell)$ is a given potential energy function. Assume that V has a second-order mininum in $\boldsymbol{\beta}_0 \in \mathcal{R}^\ell$; i.e., $\partial_{\boldsymbol{\beta}} V(\boldsymbol{\beta}) = \mathbf{0}$ in $\boldsymbol{\beta}_0$ and that the matrix

$$I_{ij} = \frac{\partial^2 V}{\partial \beta_i \partial \beta_j}(\boldsymbol{\beta}_0), \qquad i, j = 1, \ldots, \ell, \tag{4.8.2}$$

is positive definite. Then $\boldsymbol{\beta}_0$ is an equilibrium point.

7 Definition. *The "small oscillations" near $\boldsymbol{\beta}_0$ of the system described by Eq. (4.8.1), with V verifying Eq. (4.8.2), are the motions of the mechanical system with Lagrangian function*

$$\mathcal{L}_{small}(\dot{\boldsymbol{\beta}}, \boldsymbol{\beta}) = \frac{1}{2} \sum_{i,j=1}^{\ell} g_{ij}(\boldsymbol{\beta}_0) \dot{\beta}_i \dot{\beta}_j - \frac{1}{2} \sum_{i,j=1}^{\ell} I_{ij}(\beta_i - \beta_{0i})(\beta_j - \beta_{0j}) \tag{4.8.3}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_\ell)$, $\boldsymbol{\beta}_0 = (\beta_{01}, \ldots, \beta_{0\ell})$. The normal modes pulsations of Eq. (4.8.3) are called the "proper pulsations" of Eq. (4.8.1) near $\boldsymbol{\beta}_0$; their reciprocals multiplied by 2π are the "proper periods". The reciprocals of the periods are the "proper frequencies" of the small oscillations.

Observations.
(1) Therefore, the small oscillations are the motions of the Lagrangian system obtained by replacing the matrix g with its value at the equilibrium point $\boldsymbol{\beta}_0$ and by replacing the potential energy V by its Taylor expansion about $\boldsymbol{\beta}_0$ truncated to second order:

$$V(\boldsymbol{\beta}) = V(\boldsymbol{\beta}_0) + \frac{1}{2} \sum_{i,j=1}^{\ell} I_{ij}\,(\beta_i - \beta_{0i})(\beta_j - \beta_{0j}) \tag{4.8.4}$$

and in Eq. (4.8.3), $V(\boldsymbol{\beta}_0)$ does not appear since it does not affect the associated motions.

(2) On the basis of the above definitions, the "small oscillations" are not necessarily motions with small amplitude. However, one can expect or hope that if the energy of a motion described by Eq. (4.8.1) is just slightly above $V(\boldsymbol{\beta}_0)$ (hence, the motion takes place in the vicinity of $\boldsymbol{\beta}_0$ when it is initially there), then the motion of Eq. (4.8.3) with the same initial data approximates well the exact motion.

Since the small oscillations are, by definition, harmonic motions, hence "simple motions", one understands the interest in the following question: in what sense do the small oscillations approximate the real motions of Eq. (4.8.1) near $\boldsymbol{\beta}_0$?

In Chapter 2 we met and essentially solved this problem for systems with one degree of freedom. The generalization to systems with $\ell > 1$ degrees of freedom is, however, surprisingly difficult and interesting. In Chapter 5 we shall discuss some of its aspects. For the moment we shall only provide a definition of a class of systems behaving "as if they were linear oscillators" and we shall continue by discussing a few remarkable examples of such systems warning the reader, however, that it should not be hoped that Definition 10 to follow is a definition covering many cases.

8 Definition. *Consider a system of N point masses in $\mathcal{R}^d$ subject to ideal bilateral constraints with ℓ degrees of freedom and to a conservative force.*
We assume that the equations of motion are normal in the future as well as in the past; i.e., they admit a global solution $t \to S_t(\dot{\mathbf{x}}, \mathbf{x})$ for every initial datum $(\dot{\mathbf{x}}, \mathbf{x})$ compatible with the constraints. We shall call "space of the initial data" the set $\mathcal{S} \subset \mathcal{R}^{2Nd}$ of all the pairs $(\dot{\mathbf{x}}, \mathbf{x})$, where $\mathbf{x}$ is a constraint compatible configuration and $\dot{\mathbf{x}}$ is a constraint-compatible velocity.
We define on $\mathcal{S}$ the "time evolution flow", $(S_t)_{t\in\mathcal{R}}$, as the group of transformations mapping $(\dot{\mathbf{x}}, \mathbf{x})$ into $S_t(\dot{\mathbf{x}}, \mathbf{x})$ = (datum into which $(\dot{\mathbf{x}}, \mathbf{x})$ evolves in the time t according to the equations of motion).

Observations
(1) This generalizes the initial data space, introduced in §2.22, to constrained systems.
(2) $\mathcal{S}$ will be considered to be a surface in $\mathcal{R}^{2Nd}$. The geometric structure of $\mathcal{S}$ is very simple as expressed by the following proposition.

18 Proposition. *The surface $\mathcal{S}$ of the preceding definition is a regular surface in $\mathcal{R}^{2Nd}$.*

Proof. By Definition 10, §3.6, p.170, given $(\dot{\mathbf{x}}_0, \mathbf{x}_0) \in \mathcal{S}$, we have to find a neighborhood W of $(\dot{\mathbf{x}}_0, \mathbf{x}_0)$ on which it is possible to establish a local system of regular coordinates adapted to the surface $\mathcal{S}$.

Let U be a neighborhood of $\mathbf{x}_0$ on which it is possible to establish a local system of regular coordinates $\boldsymbol{\xi} = \boldsymbol{\Xi}(\boldsymbol{\beta})$, with basis Ω, adapted to the surface Σ in $\mathcal{R}^{Nd}$ defined by the constraint. The set U exists by the very definition of an ℓ-degrees-of-freedom holonomous constraint.

In this coordinate system, the possible velocity vectors $\dot{\boldsymbol{\beta}}$ for the system which are compatible with the constraints are those such that

$$\dot{\beta}_1 = \dot{\beta}_2 = \ldots = \dot{\beta}_{Nd-\ell} = 0, \tag{4.8.5}$$

while the possible position vectors $\boldsymbol{\beta}$ are those for which

$$\beta_1 = \beta_2 = \ldots = \beta_{Nd-\ell} = 0, \qquad (0, \ldots, 0, \beta_{Nd-\ell+1}, \ldots, \beta_{Nd}) \in \Omega. \tag{4.8.6}$$

Hence, the correspondence between $R^{Nd} \times \Omega$ and $\mathcal{R}^{2Nd}$ described by

$$\dot{\mathbf{x}}^{(i)} = \sum_{k=1}^{Nd} \dot{\beta}_k \frac{\partial \boldsymbol{\Xi}^{(i)}}{\partial \beta_k}(\boldsymbol{\beta}), \qquad \mathbf{x}^{(i)} = \boldsymbol{\Xi}(\boldsymbol{\beta}) \tag{4.8.7}$$

establishes on the image $W \subset \mathcal{R}^{2Nd}$ of $\mathcal{R}^{Nd} \times \Omega$ a coordinate system near $(\dot{\mathbf{x}}_0, \mathbf{x}_0) \in W$ adapted to $\mathcal{S}$ with basis $\mathcal{R}^{Nd} \times \Omega$, and it is easily checked that the Jacobian determinant of this coordinate change at the point with coordinates $(\dot{\boldsymbol{\beta}}, \boldsymbol{\beta})$ is the square of the Jacobian determinant in $\boldsymbol{\beta}$ of the transformation $\boldsymbol{\Xi}$. By the regularity assumption, on the coordinate system $(U, \boldsymbol{\Xi})$, such a determinant does not vanish. mbe

Observations.
(1) The above proof shows that setting

$$\begin{aligned} \dot{\boldsymbol{\kappa}} &= (\dot{\beta}_{Nd-\ell+1}, \ldots, \dot{\beta}_{Nd}) \stackrel{def}{=} (\dot{\kappa}_1, \ldots, \dot{\kappa}_\ell), \\ \boldsymbol{\kappa} &= (\beta_{Nd-\ell+1}, \ldots, \beta_{Nd}) \stackrel{def}{=} (\kappa_1, \ldots, \kappa_\ell), \end{aligned} \tag{4.8.8}$$

Eqs. (4.8.7) establish a coordinate system, $(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$, for the points of $W \cap \mathcal{S}$, where W is the image via Eqs. (4.8.7) of $\mathcal{R}^{Nd} \times \Omega$. Furthermore, as $(\dot{\mathbf{x}}, \mathbf{x})$ varies in $W \cap \mathcal{S}$, the point $(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$ varies in $\mathcal{R}^\ell \times V$ where V is an open convex set in $\mathcal{R}^d$ (as $V = \Omega \cap \{\text{plane } \beta_1, \ldots, \beta_{Nd-\ell} = 0\}$, and Ω is convex).
One refers to this remark by saying that the data space $\mathcal{S}$ of a system with ℓ degrees of freedom locally has the structure $\mathcal{R}^\ell \times V$ with $V \subset \mathcal{R}^\ell$".
For this reason, and with an abuse of notation very useful and widely used, one often denotes the points of $\mathcal{S}$ as $(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$, where $(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$ are local regular coordinates (which have to be deduced from the context and which often are really local (i.e., non global) coordinates), in a neighborhood W of a point in $\mathcal{S}$ such that $W \cap \mathcal{S}$ has the structure $\mathcal{R}^\ell \times V$.
Coherently, the Lagrangian of the constrained system is described as a function $\mathcal{L}(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$ of $(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$.
(2) Since $\mathcal{S}$ is a regular surface, it makes sense to define the open sets on $\mathcal{S}$ and the space $C^\infty(\mathcal{S})$. A set $E \in \mathcal{S}$ is open on $\mathcal{S}$ if it is the intersection of an open set in $\mathcal{R}^{Nd}$ with $\mathcal{S}$. A function f is in $C^\infty(\mathcal{S})$ if its restriction to a neighborhood U, on which it is possible to set up a local system of regular coordinates transforming U into $\mathcal{R}^\ell \times V$, has the property that, if thought of as a function of the local coordinates $(\dot{\boldsymbol{\kappa}}, \boldsymbol{\kappa})$, it is a $C^\infty(\mathcal{R}^\ell \times V)$ function.

9 Definition. *Let $\mathcal{S}$ be the initial data space for a system of N point masses with ℓ degrees of freedom subject to ideal holonomous constraints and to conservative forces.Let $A \in C^\infty(\mathcal{S})$ be a real-valued function on $\mathcal{S}$. We shall say that A is a "prime integral" (or "first integral" or "constant of motion") for the motions $t \to S_t(\dot{\mathbf{x}}, \mathbf{x}) \equiv (\dot{\mathbf{x}}(t), \mathbf{x}(t))$, $t \in \mathcal{R}$, if*

$$A(\dot{\mathbf{x}}(t), \mathbf{x}(t)) = \text{constant} \tag{4.8.9}$$

for all $(\dot{\mathbf{x}}, \mathbf{x}) \in \mathcal{S}$.

Examples

(1) The energy

$$E(\dot{\mathbf{x}}, \mathbf{x}) = \frac{1}{2}\sum_{i=1}^{N} m_i(\dot{\mathbf{x}}^{(i)})^2 + V^{(a)}(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}) \tag{4.8.10}$$

is a typical example of a prime integral. Often it is the only prime integral admitted by the system's motions.

(2) If the system is isolated, i.e., subject to zero external forces, the d components of the linear momentum

$$\mathbf{Q}(\dot{\mathbf{x}}, \mathbf{x}) = \sum_{i=1}^{N} m_i\, \dot{x}^{(i)} \tag{4.8.11}$$

are also prime integrals when the third law of dynamics holds. In the same situation, the angular momentum components also give rise to prime integrals.

19 Proposition. *A system of ℓ harmonic oscillators with Lagrangian function (4.1.2) admits ℓ prime integrals $A_1, \ldots, A_\ell$ given by Eq. (4.1.5). Furthermore, it is possible to parameterize the initial data space $\mathcal{S}$ through the values $A_1, \ldots, A_\ell$ and a point $\boldsymbol{\varphi} \in \mathcal{T}^\ell$, $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_\ell)$, so that $\mathcal{S}$ can be thought of as the product $[0, +\infty) \times \mathcal{T}^\ell$, and the motion $t \to (\dot{\mathbf{x}}(t), \mathbf{x}(t))$, $t \in \mathcal{R}_+$, of the system is described, in these coordinates, as*

$$(A_1, \ldots, A_\ell; \varphi_1, \ldots, \varphi_\ell) \to (A_1, \ldots, A_\ell; \varphi_1 + \omega_1 t, \ldots, \varphi_\ell \omega_\ell t), \tag{4.8.12}$$

where $\omega_1, \ldots, \omega_\ell$ are positive constants.

Finally, the correspondence $(\mathbf{A}, \boldsymbol{\varphi}) \to (\dot{\mathbf{x}}, \mathbf{x})$ is a C^∞ invertible nonsingular[6] *correspondence between $(0, +\infty)^\ell \times \mathcal{T}^\ell$ and the subset of $\mathcal{R}^{2\ell}$ which is its image.*

Proposition 19 suggests the following definition.

10 Definition. *Let $\mathcal{S}$ be the initial data space for a system with ℓ degrees of freedom subject to ideal constraints and to conservative active forces.*

We shall say that the system is "integrable" on the open region $W \subset \mathcal{S}$ if on W it is possible to define ℓ prime integrals $\mathbf{A} = (A_1, \ldots, A_\ell)$ and ℓ $\mathcal{T}^\ell$-valued $C^\infty(W)$ functions $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_\ell)$ such that:

(1) The image of W under the map

[6] See definition 13 and related observations, p.101, for the meaning of the derivatives.

$$(\dot{\mathbf{x}}, \mathbf{x}) \to I(\dot{\mathbf{x}}, \mathbf{x}) = (\mathbf{A}, \boldsymbol{\varphi}) \tag{4.8.13}$$

has the form $V \times \mathcal{T}^\ell$, where V is an open set in $\mathcal{R}^\ell$ and the correspondence I between W and $V \times \mathcal{T}^\ell$ is an invertible nonsingular (i.e., with non vanishing Jacobian determinant) correspondence.
(2) There are ℓ real C^∞ functions on V, $\mathbf{A} \to \boldsymbol{\omega}(\mathbf{A}) = (\omega_1(\mathbf{A}), \ldots, \omega_\ell(\mathbf{A})) \in \mathcal{R}^\ell$ such that if $t \to \mathbf{x}(t)$ is a motion with initial data $(\dot{\mathbf{x}}(0), \mathbf{x}(0)) \in W$, then, $\forall\, t \in \mathcal{R}_+$, $(\dot{\mathbf{x}}(t), \mathbf{x}(t)) \in W$ and

$$I(\dot{\mathbf{x}}(t), \mathbf{x}(t)) = (\mathbf{A}_0, \boldsymbol{\varphi}_0 + \boldsymbol{\omega}(\mathbf{A}_0)t) \tag{4.8.14}$$

where $\mathbf{A}_0 = \mathbf{A}(\dot{\mathbf{x}}(0), \mathbf{x}(0))$, $\boldsymbol{\varphi}_0 = \boldsymbol{\varphi}(\dot{\mathbf{x}}(0), \mathbf{x}(0))$, and $\boldsymbol{\varphi} \to \boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{A}_0)t$ denotes the quasi-periodic flow on $\mathcal{T}^\ell$ with speed $\boldsymbol{\omega}(\mathbf{A}_0)$, see Definition 1, p.248.. The numbers $\omega_i(\mathbf{A}), T_i(\mathbf{A}) = \frac{2\pi}{\omega_i(\mathbf{A})}$, $\nu_i(\mathbf{A}) = \frac{1}{T_i(\mathbf{A})}$, $i = 1, \ldots, \ell$, are, respectively, called the pulsations, the periods, and the frequencies of the motions in W with amplitudes $\mathbf{A}$.

Observations. (1) In the case of a system of harmonic oscillators, there are various choices of W for which the system is integrable on W: the most natural one takes W to be the set in $\mathcal{S}$ whose image under the map of Eq. (4.1.15) is $(0, +\infty)^\ell \times \mathcal{T}^\ell$ (i.e., the set of data having all the normal modes excited: $A_i > 0$ for $i = 1, \ldots, \ell$).
(2) One can interpret Eqs. (4.8.13) and (4.8.14) as saying that the data space W of an integrable system is "foliated by an ℓ-parameter family of ℓ-dimensional invariant tori". The parameters are the values of the ℓ prime integrals. The torus with parameters $\mathbf{A} \in V$ is the set $I(\{\mathbf{A}\} \times \mathcal{T}^\ell)$ image of $\{\mathbf{A}\} \times \mathcal{T}^\ell$ under the "integration map" I.
(3) In the case of harmonic oscillators, $\boldsymbol{\omega}(\mathbf{A})$ is $\mathbf{A}$ independent: "isochrony of the harmonic oscillations". As seen in the case $\ell = 1$, §2.10, it is obvious that this should be a very special property of the harmonic oscillators. Therefore, it is better not to introduce it into the definition of integrable system, to avoid giving a too restrictive definition.
(4) In the context of the theory of small oscillations, the above definition seems especially designed to formulate the conjecture that in a small enough neighborhood W of an equilibrium position $(\mathbf{0}, \boldsymbol{\beta}_0)$ for a mechanical system described by a Lagrangian (4.8.1) verifying Eq. (4.8.2), the system is integrable.
Such a conjecture, true if $\ell = 1$, is generally false if $\ell > 1$; i.e., there may be motions which stay indefinitely close to an equilibrium point and, nevertheless, move in a fashion substantially different from a quasi-periodic motion. However, a conjecture similar to this one is true. We shall discuss this matter in Chapter 5, §5.9-§5.12.
(5) To establish the integrability of a system with ℓ degrees of freedom, one usually proceeds to show that it is possible to describe the motions which develop in W in terms of 2ℓ parameters $(\mathbf{A}, \boldsymbol{\varphi}) \in V \times \mathcal{T}^\ell$ and of N C^∞-

functions on $V \times \mathcal{T}^\ell$, $\boldsymbol{\Phi}^{(1)}, \ldots, \boldsymbol{\Phi}^{(N)}$, such that if $t \to \mathbf{x}(t)$ is a motion, one has, $\forall\, i = 1, 2, \ldots, N$,

$$\mathbf{x}^{(i)}(t) = \boldsymbol{\Phi}^{(i)}(A_1, \ldots, A_\ell, \varphi_1 + \omega_1\, t, \ldots, \varphi_\ell + \omega_\ell\, t), \tag{4.8.15}$$

where $\omega_1(\mathbf{A}), \ldots, \omega_\ell(\mathbf{A})$ are ℓ C^∞-functions of $\mathbf{A} \in V$.
Successively, one proceeds to check the invertibility, regularity, and non singularity of the map $\dot{\mathbf{x}}(0), \mathbf{x}(0)) \longleftrightarrow (\mathbf{A}, \boldsymbol{\varphi})$. This check is usually an easy matter and without direct interest once Eq. (4.8.15) has been established for all the motions in W. Actually, the true analytic difficulty that is met in the intregrability proofs lies in the proof of the validity of a consequence of Eq. (4.8.15): precisely, in checking that all the motions in W are "quasi periodic" in the sense that their coordinates depend quasi-periodically on time (see §2.21 for the notion of quasi-periodicity). See, however, Problem 20 to §4.15.

Therefore, in the upcoming sections, we shall often stop our analysis of integrability when we find that the motions taking place in a given W are quasi-periodic, without entering into the sometimes long analysis necessary to prove the invertibility and smoothness properties required by integrability.

The following extension of Definition 10 is natural in the context of the concepts of analytical mechanics of §3.11 and §3.12.

11 Definition. *Let $\mathcal{L} \in C^\infty(W), W \subset \mathcal{R}^{2\ell}$ or $W \subset (\mathcal{R}^\ell \times \mathcal{T}^\ell)$ or $W \subset \mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2})$, $\ell_1 + \ell_2 = \ell$ be a time-independent regular Lagrangian on W (see Definition 14, §3.11, p.211).*
We say that $\mathcal{L}$ is integrable on the data space W if there is an integrating map I transforming W into $V \times \mathcal{T}^\ell$ enjoying the properties (1) and (2) of Definition 10, where the motion $t \to \mathbf{x}(t)$ is now a solution to the Lagrangian equations relative to $\mathcal{L}$.
Similarly, if $H \in C^\infty(\widetilde{W})$, $\widetilde{W} \subset \mathcal{R}^{2\ell}$ or $\widetilde{W} \subset \mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\widetilde{W} \subset \mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2})$, $\ell_1 + \ell_2 = \ell$, is a regular time-independent Hamiltonian function on the phase space $\widetilde{W}$, we say that H is integrable on $\widetilde{W}$ if the corresponding Lagrangian function $\mathcal{L}$ is integrable on the data space subset $W = \boldsymbol{\Xi}^{-1}(\widetilde{W})$, $\boldsymbol{\Xi}$ being the map inducing the Legendre transformation between H and $\mathcal{L}$ (see §3.11).

In this case, if I is the integrating map for $\mathcal{L}$, the map

$$\widetilde{I}(\mathbf{p}, \mathbf{q}) = I(\boldsymbol{\Xi}^{-1}(\mathbf{p}, \mathbf{q})) \tag{4.8.16}$$

maps W onto $V \times \mathcal{T}^\ell$ and it is called an "integrating map" for H.
If $\widetilde{I}$ is a completely canonical map of $\widetilde{W}$ onto $V \times \mathcal{T}^\ell$ we say that H is "canonically integrable" on the phase space $\widetilde{W}$.
If H is analytic[7] *on $\widetilde{W}$ and $\widetilde{I}$ is also analytic, we say that H is "analytically*

[7] Analytic means "having convergent Taylor series" near every point of the domain of definition, see Definitions 13,14 and 15, §4.13, p.336.

integrable on $\widetilde{W}$".If H is analytic and if $\widetilde{I}$ is an analytic completely canonical map of W onto $V \times \mathcal{T}^\ell$, we shall say that H is "canonically analytically integrable" on $\widetilde{W}$.

Observations.
(1) If $H \in C^\infty(W)$ is canonically integrable and if$\widetilde{I}$ is a completely canonical map integrating H, then

$$H(I^{-1}(\mathbf{A}, \boldsymbol{\varphi})) \equiv h(\mathbf{A}) = (\boldsymbol{\varphi}\text{-independent function}) \tag{4.8.17}$$

and $\boldsymbol{\omega}(\mathbf{A}) = (\frac{\partial h}{\partial A_1}(\mathbf{A}), \ldots, \frac{\partial h}{\partial A_\ell}(\mathbf{A}))$:

$$\boldsymbol{\omega}(\mathbf{A}) = \frac{\partial h}{\partial \mathbf{A}}(\mathbf{A}) \tag{4.8.18}$$

In this case the variables $(\mathbf{A}, \boldsymbol{\varphi})$ are called "action-angle" variables and are canonical variables.
(2) It turns out that all the systems that we shall consider in the upcoming sections are analytically canonically integrable on vast regions of phase space. However, this will not always be explicitly checked and it will be left to the reader, in the problems, to draw this conclusion from the properties discussed in the text.
(3) In an obvious way, one could also define the notion of a Lagrangian analytically integrable on some set W in the data space. The corresponding Hamiltonian system would then be analytically integrable on the corresponding phase-space subset $\widetilde{W}$ and vice versa.

4.8.1 Problems

1. Given $\boldsymbol{\omega} \in \mathcal{R}^\ell$ and $g \in C^\infty(\mathcal{T}^\ell)$, suppose that $\forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell$, $\boldsymbol{\nu} \neq \mathbf{0}$, it is $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|^{-1} < C|\boldsymbol{\nu}|^\alpha$, for some $C > 0, \alpha > 0$. Show that the system on $\mathcal{R}^\ell \times \mathcal{T}^\ell$ with Hamiltonian $(\mathbf{A}, \boldsymbol{\varphi}) = \mathbf{A}\cdot\boldsymbol{\omega} + g(\boldsymbol{\varphi})$ is integrable and find an expression for ℓ prime integrals. Show that this is an isochronous system. Note that the equations of motion can be solved explicitly for general $\boldsymbol{\omega}$. (*Hint:* Write the equations of motion and solve the one for the $\mathbf{A}$'s by developing g into a Fourier series $g(\boldsymbol{\varphi}) = \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell} \widehat{g}_{\boldsymbol{\nu}} e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}$ before integration. The prime integrals can be chosen

$$\mathbf{B} = \mathbf{A} + \sum_{\substack{\boldsymbol{\nu} \neq \mathbf{0} \\ \boldsymbol{\nu} \in \mathcal{Z}^\ell}} \boldsymbol{\nu}\, \widehat{g}_{\boldsymbol{\nu}} \frac{e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{\boldsymbol{\nu}\cdot\boldsymbol{\omega}}$$

and the condition on $\boldsymbol{\omega}$ is required to insure the convergence of the series.)

2. In the context of Problem 1, show that if g is a trigonometric polynomial (i.e., it has finitely many non vanishing Fourier coefficients), then the results of Problem 1 hold under the sole assumption that the components of $\boldsymbol{\omega}$ are rationally independent.

3. In the context of Problem l, suppose that there is $\boldsymbol{\nu}_0 \in \mathcal{Z}^\ell, \boldsymbol{\nu}_0 \neq \mathbf{0}$, such that $\boldsymbol{\omega}\cdot\boldsymbol{\nu}_0 = 0$. Show that the Hamiltonian system with Hamiltonian $H(\mathbf{A}, \boldsymbol{\varphi}) = \mathbf{A}\cdot\boldsymbol{\omega} + \varepsilon cos(\boldsymbol{\nu}_0\cdot\boldsymbol{\varphi})$ is not integrable. (*Hint:* Show that its motions are not quasi-periodic.)

4. In the context of Problem 1, suppose that $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|^{-1} < C|\boldsymbol{\nu}|^\alpha, \forall\, \boldsymbol{\nu} \neq \mathbf{0}$ and $\boldsymbol{\nu}_0 \notin N_0$, where N_0 is a subset of $\mathcal{Z}^\ell$. Suppose also that $g \in C^\infty(\mathcal{T}^\ell)$ is such that $\widehat{g}_{\boldsymbol{\nu}} = 0$, $\forall\, \boldsymbol{\nu} \in N_0$. Show that the Hamiltonian $H(\mathbf{A}, \boldsymbol{\varphi}) = \mathbf{A}\cdot\boldsymbol{\omega} + g(\boldsymbol{\varphi})$ is integrable on $\mathcal{R} \times \mathcal{T}^\ell$.

5. Show that the integrability in Problems 1, 2, and 4 is analytical and canonical (*Hint:* $\mathbf{A}'\cdot\boldsymbol{\varphi}+\sum_{\boldsymbol{\nu}\neq\mathbf{0}}\widehat{g}_{\boldsymbol{\nu}}\,\frac{e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i\,\boldsymbol{\omega}\cdot\boldsymbol{\nu}}\stackrel{def}{=}\Phi(\mathbf{A}',\boldsymbol{\varphi})$ is a generating function for the integrating map.)

4.9 Integrable Systems. Central Motions with Non vanishing Areal Velocity. The Two-Body Problem

The best-known integrable mechanical system consists, perhaps, of two point masses with masses $m_1, m_2 > 0$ interacting through a conservative force with potential energy $\overline{V}$ depending only on the distance between the two points:

$$\overline{V}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)=V(|\boldsymbol{\xi}_1-\boldsymbol{\xi}_2|); \tag{4.9.1}$$

and we shall assume that the function $\varrho\to V(\varrho)$ is defined for $\varrho>0$ and that it is a C^∞-function such that

$$\lim_{\varrho\to 0}\varrho^2 V(\varrho)=0,\qquad \inf_{\varrho>\varepsilon}V(\varrho)=-V_\varepsilon,\quad \forall\,\varepsilon>0 \tag{4.9.2}$$

Note that $V(0)$ is undefined, and this means that we shall only consider motions $t\to(\mathbf{x}_1(t),\mathbf{x}_2(t))$, $t\in\mathcal{R}_+$, such that $|\mathbf{x}_1(t)-\mathbf{x}_2(t)|>0$, $t\in\mathcal{R}_+$. This restriction will be imposed via the condition of non vanishing areal velocity.

If $t\to(\mathbf{x}_1(t),\mathbf{x}_2(t))$, $t\in\mathcal{R}_+$, is a motion of the system, the two points will move so that the total linear and angular momentum will be conserved. In fact, the force generated by Eq. (4.9.1) is easily seen to verify the third law of dynamics, so that the cardinal equations hold and imply the above mentioned conservation laws.

Hence, the center of mass G moves in a uniform rectilinear fashion and, possibly by changing reference system, it may be supposed that G coincides with the origin O of the reference system $(O;\mathbf{i},\mathbf{j},\mathbf{k})$ in which motion is studied. In this situation, the motion $t\to(\mathbf{x}_1(t),\mathbf{x}_2(t))$, $t\in\mathcal{R}_+$ will be such that

$$m_1\mathbf{x}_1(t)=-m_2\mathbf{x}_2(t),\qquad \forall\,t\in\mathcal{R}_+ \tag{4.9.3}$$

and to determine the positions of the two points it will suffice to give the vector $\boldsymbol{\varrho}\stackrel{def}{=}\mathbf{x}_2(t)-\mathbf{x}_1(t)$:

$$\mathbf{x}_1(t)=-\frac{m_2}{m_1+m_2}\boldsymbol{\varrho}(t),\qquad \mathbf{x}_2(t)=\frac{m_1}{m_1+m_2}\boldsymbol{\varrho}(t), \tag{4.9.4}$$

Since the angular momentum with respect to O is a constant vector $\mathbf{K}$, it can be assumed, without loss of generality, that $\mathbf{K}$ is parallel to $\mathbf{k}$:

$$\mathbf{K}=\widetilde{A}\,\mathbf{k}. \tag{4.9.5}$$

Only motions for which $\widetilde{A}>0$ will be considered and it will be seen that $\widetilde{A}$ is proportional to the areal velocity. From Eqs. (4.9.3)-(4.9.5), it follows that

$$\mathbf{K} = \widetilde{A}\mathbf{k} = m_1\mathbf{x}_1 \wedge \dot{\mathbf{x}}_1 + m_2\mathbf{x}_2 \wedge \dot{\mathbf{x}}_2 = m_1\mathbf{x}_1 \wedge (\dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2) = -\frac{m_1 m_2}{m_1 + m_2}\boldsymbol{\varrho} \wedge \dot{\boldsymbol{\varrho}}. \tag{4.9.6}$$

i.e., $\boldsymbol{\varrho}$ and $\dot{\boldsymbol{\varrho}}$ must both lie in the plane $(\mathbf{i}, \mathbf{j})$. Therefore, the motion $t \to \boldsymbol{\varrho}(t)$, $t \in \mathcal{R}_+$, takes place on the plane $(\mathbf{i}, \mathbf{j})$. Recalling the considerations in §3.4 about the constraints, we can find the equations of motion by parameterizing the motion by the polar coordinates (ϱ, θ) of $\boldsymbol{\varrho}$ in the plane $(\mathbf{i}, \mathbf{j})$and then writing the Lagrangian equations for the Lagrangian

$$\mathcal{L} = \frac{1}{2}m_1\dot{\mathbf{x}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{x}}_2^2 - V(|\mathbf{x}_1 - \mathbf{x}_2|) \tag{4.9.7}$$

computed on the motions, parameterized as above.

For such motions, Eq. (4.9.7) becomes

$$\begin{aligned} &\mathcal{L}(\dot{\varrho}, \dot{\theta}, \varrho, \theta) = \frac{m_1}{2}\big(\frac{m_2}{m_1 + m_2}\big)^2 \dot{\boldsymbol{\varrho}}^2 + \frac{m_2}{2}\big(\frac{m_1}{m_1 + m_2}\big)^2 \dot{\boldsymbol{\varrho}}^2 - V(\varrho) \\ &= \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}\dot{\boldsymbol{\varrho}}^2 - V(\varrho) = \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}(\dot{\varrho}^2 + \dot{\varrho}^2\dot{\theta}^2) - V(\varrho), \end{aligned} \tag{4.9.8}$$

where the well-known formula expressing the square velocity $\dot{\boldsymbol{\varrho}}^2$ as $\dot{\varrho}^2 + \varrho^2\dot{\theta}^2$ in polar coordinates has been used together with Eq. (4.9.4). Equation (4.9.8) yields the following proposition:

20 Proposition. *The theory of the motion of two point masses, with masses $m_1, m_2 > 0$, under the action of a mutual central conservative force with potential energy given by Eq. (4.9.1) is equivalent to the theory of the motion of a single point mass with mass m:*

$$m = \frac{m_1 m_2}{m_1 + m_2} \tag{4.9.9}$$

moving on a plane under the action of a conservative force, centrally acting on the mass from a point O in the plane, with the same potential energy V.

The motions described by the Lagrangian function (4.9.8) and such that $\widetilde{A} \neq 0$ are called "central motions".

21 Proposition. *The motions of the mechanical system described by Eq. (4.9.8) admit two prime integrals:*

$$E = \frac{1}{2}m\dot{\varrho}^2 + \varrho^2\dot{\theta}^2 + V(\varrho), \tag{4.9.10}$$

$$A = \varrho^2\dot{\theta}, \tag{4.9.11}$$

and, if $A \neq 0$, they indefinitely stay away from the origin at a distance greater than some time-independent positive quantity (A and E independent).

Proof. Equation (4.9.10) is the total energy and, by Eq. (4.9.6), $\widetilde{A} = m\varrho^2\dot\theta$ the angular momentum along the z axis. Hence, E and A are both prime integrals. Note that $\varrho^2\dot\theta$ is twice the "areal velocity", i.e., twice the area spanned by $\boldsymbol{\varrho}$ per unit time.

By substituting Eq. (4.9.11) into Eq. (4.9.10) it follows

$$E = \frac{1}{2}m(\dot\varrho^2 + \varrho^2\dot\theta^2) + V(\varrho) \tag{4.9.12}$$

and Eq. (4.9.2) implies the existence of $\varrho_0 > 0$ such that $E - m\frac{A^2}{2\varrho^2} - V(p) < 0$ for $\varrho < \varrho_0$. So $\varrho(t) > \varrho_0$, $\forall\, t \in \mathcal{R}_+$. mbe

Let $t \to (\varrho(t), \theta(t))$, $t \in \mathcal{R}_+$ be a motion associated with Eq. (4.9.8) with $A > 0$. Write the equation of motion for ϱ by considering the Lagrangian equation relative to Eq. (4.9.8) and corresponding to the coordinate ϱ:

$$m\ddot\varrho = m\varrho\dot\theta^2 - \frac{\partial V}{\partial \varrho}. \tag{4.9.13}$$

By Eq. (4.9.11) the latter relation becomes

$$m\ddot\varrho = m\frac{A^2}{\varrho^3} - \frac{\partial V}{\partial \varrho}(\varrho) \equiv -\frac{\partial V_A}{\partial \varrho}(\varrho) \tag{4.9.14}$$

where

$$V_A(\varrho) = \frac{mA^2}{2\varrho^2} + V(\varrho) \tag{4.9.15}$$

showing that the ϱ coordinate evolves in time as the abscissa of a mass m on a line, subject to a conservative force with potential energy V_A.

Since the motion, by Proposition 21, is such that $\varrho(t) \geq \varrho_0 > 0$, we can ignore the singularities of V and V_A in $\varrho = 0$ and we can also ignore the constraint $\varrho > 0$ due to ϱ being the polar radial coordinate, so that the theory of Chapter 2 for conservative C^∞ forces acting upon one-dimensional systems.

22 Proposition. *Let $\varrho \to V(\varrho)$, $\varrho > 0$, be a $C^\infty((0, +\infty))$ function verifying Eq. (4.9.2). Let W be the open set, in the data space of the system described by Eq. (4.9.8), consisting of the data with E and A in Eqs. (4.9.10) and (4.9.11) such that*

$$(i) \qquad A > 0. \tag{4.9.16}$$

$$(ii)\ \textit{The equation} \qquad V_A(\varrho) = \frac{mA^2}{2\varrho^2} + V(\varrho) = E \tag{4.9.17}$$

admits just two solutions $\varrho_-(E, A), \varrho_+(E, A)$ such that $\varrho_+(E, A) > \varrho_-(E, A)$ and $\frac{\partial V_A}{\partial \varrho}(\varrho_\pm(E, A)) \neq 0$.
Then the system is integrable in a neighborhood of every point in W and has two periods, see Definition 10, §4.8, p.287, given by

$$T_1(E,A) = 2\int_{\varrho_-}^{\varrho_+} \frac{d\varrho}{\sqrt{\frac{2}{m}(E - V_A(\varrho))}}, \tag{4.9.18}$$

$$T_2(E,A) = \frac{2\pi}{A} \frac{\int_{\varrho_-}^{\varrho_+} \frac{d\varrho}{\sqrt{\frac{2}{m}(E-V_A(\varrho))}}}{\int_{\varrho_-}^{\varrho_+} \frac{d\varrho}{\varrho^2\sqrt{\frac{2}{m}(E-V_A(\varrho))}}}, \tag{4.9.19}$$

where $\varrho_+ = \varrho_+(E,A)$ *and* $\varrho_- = \varrho_-(E,A)$.

PROOF. In the course of the proof we shall state that some functions are C^∞, leaving the proof to the reader. Let $(\dot\varrho_0, \dot\theta_0, \varrho_0, \theta_0) \in W$ be an initial datum with energy E and areal velocity $\frac{A}{2}$ and consider the solution of Eq. (4.9.14),

$$t \to R(t,E,A), \qquad t \in \mathcal{R}_+ \tag{4.9.20}$$

with initial datum

$$R(0,E,A) = \varrho_-(E,A), \qquad \dot R(0,E,A) = 0. \tag{4.9.21}$$

By the theory of one-dimensional motions, §2.7, the function R is a C^∞ function periodic in t with period

$$T_1(E,A) = 2\int_{\varrho_-}^{\varrho_+} \frac{d\varrho}{\sqrt{\frac{2}{m}(E - V_A(\varrho))}}, \tag{4.9.22}$$

where $\varrho_\pm = \varrho_\pm(E,A)$

If $t_0(\varrho_0, \dot\varrho_0)$ is the shortest time such that

$$R(t_0,E,A) = \varrho_0, \qquad \dot R(t_0,E,A) = \dot\varrho_0, \tag{4.9.23}$$

necessarily existing by our assumptions on W, it follows that

$$\varrho(t) = R(t + t_0(\varrho_0,\dot\varrho_0), E, A), \qquad t \in \mathcal{R}_+. \tag{4.9.24}$$

To complete the analysis of the motion, it is necessary to determine $\theta(t)$. Using Eq. (4.9.11):

$$\theta(t) = \theta_0 + \int_0^t \frac{A}{R(t' + t_0(\varrho_0,\dot\varrho_0),E,A)^2}\, dt'; \tag{4.9.25}$$

and remark that the integrand function in Eq. (4.9.25) is a C^∞ periodic function of t' with the period of Eq. (4.9.22), since such is R and also $R \geq \varrho_-(E,A) > 0$. Then by the Fourier theorem, if $T_1 \equiv T_1(E,A)$,

$$\frac{A}{R(t,E,A)^2} = \sum_{k\in\mathcal{Z}} \chi_k(A,E) e^{\frac{2\pi}{T_1}kt}, \tag{4.9.26}$$

where $(\chi_k)_{k\in\mathcal{Z}}$ are the Fourier coefficients of $\frac{A}{R^2}$. They vanish as $k \to \infty$ faster than any power in k. Inserting Eq. (4.9.26) into Eq. (4.9.25), it appears that

$$\theta(t) = \theta_0 + \chi_0(A,E)t + \sum_{\substack{k\in\mathcal{Z}\\ k\neq 0}} \chi_k(A,E)\frac{e^{\frac{2\pi i}{T_1}kt}-1}{\frac{2\pi i k}{T_1}} e^{\frac{2\pi i}{T_1}kt_0(\varrho_0,\dot\varrho_0)} \tag{4.9.27}$$

which we shall write as

$$\theta(t) = \theta_0 + \chi_0(A,E)t + S(t+t_0(\varrho_0,\dot\varrho_0),E,A) - S(t_0(\varrho_0,\dot\varrho_0),E,A) \tag{4.9.28}$$

where

$$S(t,E,A) \stackrel{def}{=} \sum_{\substack{k\in\mathcal{Z}\\ k\neq 0}} \chi_k(A,E)\frac{e^{\frac{2\pi i}{T_1}kt}}{\frac{2\pi i k}{T_1}} \tag{4.9.29}$$

is a C^∞-function, periodic with period $T_1 \equiv T_1(E,A)$.

It is then clear that the coordinates of $\varrho(t)$ have the form of Eq. (4.8.15). For instance, if $\boldsymbol{\varrho}(t) = (\varrho_1(t), \varrho_2(t) \in \mathcal{R}^2$:

$$\begin{aligned}\varrho_1(t) =& \varrho(t)\cos\theta(t) = R(t+t_0)\,\cos(\theta_0+\chi_0 t + S(t+t_0) - S(t_0))\\ =& R(t+t_0)\Big(\cos(\theta_0+\chi_0 t)\cos(S(t+t_0)-S(t_0))\\ &-\sin(\theta_0+\chi_0 t)\sin(S(t+t_0)-S(t_0))\Big),\end{aligned} \tag{4.9.30}$$

where the dependence on the $E, A, \varrho_0, \dot\varrho_0$ variables has not been explicitly written. By Observation 4 to Definition 10, p.288, this shows the integrability of the system and that the two periods are $T_1(E,A)$ and $T_2(E,A) = \frac{2\pi}{P}\chi_0(A,E)$.

It is also easy to find explicitly the integrating transformation I: the prime integrals are E and A, the angles $(\varphi_1,\varphi_2) \in \mathcal{T}^2$ are, for instance, by Eqs. (4.9.24) and (4.9.28),

$$\varphi_1(\dot\varrho_0,\dot\theta_0,\varrho_0,\theta_0) = \frac{2\pi}{T_1(E,A)}t_0(\varrho_0,\dot\varrho_0), \qquad (\text{"average anomaly"}), \tag{4.9.31}$$

$$\varphi_2(\dot\varrho_0,\dot\theta_0,\varrho_0,\theta_0) = \theta_0 - S(t_0(\varrho_0,\dot\varrho_0),E,A) \quad (\text{"average longitude"}), \tag{4.9.32}$$

and the respective periods are, as already mentioned, $T_1(E,A)$ and $T_2(E,A)$ [see Eq. (4.9.28)].

Regularity and invertibility of the transformation I on suitable neighborhoods of the trajectory starting in $(\dot\varrho_0,\dot\theta_0,\varrho_0,\theta_0)$ will not be explicitly checked.

It remains to check Eq. (4.9.19). Again we do not write explicitly the E and A dependence in the functions $\varrho_-(E,A), \varrho_+(E,A), \chi_0(A,E), T_1(E,A), R(t,E,A), S(t,E,A)$. By the Fourier theorem,

$$\chi_0 = \frac{1}{T_1}\int_0^{T_1} \frac{A}{R(t)^2}dt = \frac{2}{T_1}\int_0^{\frac{T_1}{2}} \frac{A}{R(t)^2}dt \tag{4.9.33}$$

because $R(t)$ behaves specularly when t varies from 0 to $\frac{T_1}{2}$ or from $\frac{T_1}{2}$ to T_1, (i.e. when R varies between ϱ_- and ϱ_+ or between ϱ_+ and ϱ_-). But for $t \in [0, \frac{T_1}{2}]$,

$$t = \int_{\varrho_-}^{R(t)} \frac{d\varrho}{\sqrt{\frac{2}{m}(E - V_A(\varrho))}} \tag{4.9.34}$$

by Eqs. (4.9.15), (4.9.20). Hence, changing variables "$t \to R$", via Eq. (4.9.34), it follows that

$$dt = \frac{dR}{\sqrt{\frac{2}{m}(E - V_A(R))}}, \tag{4.9.35}$$

and this implies, from Eq. (4.9.33), that

$$\chi_0 = \frac{2A}{T_1}\int_{\varrho_-}^{\varrho_+} \frac{dR}{R^2\sqrt{\frac{2}{m}(E - V_A(R))}}, \tag{4.9.36}$$

mbe

Observation. If we regard Eq. (4.9.8) as defining a three-dimensional problem with Lagrangian

$$\mathcal{L}(\dot{\boldsymbol{\varrho}}, \boldsymbol{\varrho}) = \frac{1}{2}m\dot{\boldsymbol{\varrho}}^2 - V(\varrho) \tag{4.9.37}$$

it follows, of course, that under the same assumptions as in Proposition 22, the system is integrable. Now the prime integrals will be E, A and the angle of inclination i of the orbital plane with the reference $(\mathbf{i}, \mathbf{j})$ plane. The third angle will be the longitude in the $(\mathbf{i}, \mathbf{j})$ plane, counted from the $\mathbf{i}$ axis (say), of the intersection of the orbital plane with the $(\mathbf{i}, \mathbf{j})$ plane ("nodes line"). However, the third angle thus defined remains constant over time. This means that the pulsations in these coordinates will be $\omega_1 = \frac{2\pi}{T_1}$, $\omega_2 = \frac{2\pi}{T_2}$, $\omega_3 = 0$.

4.9.1 Problems

1. Let $m = 1$ and consider the motions associated with the Lagrangian (4.9.8) under the assumptions of Proposition 22. Following the idea of Problem 4, p.227, and substituting L for A in that problem, define

$$L \stackrel{def}{=} \lambda(E, A) = \int_{\varrho_-(E,)}^{\varrho_+(E,)} \sqrt{2(E - V_A(\varrho))}\,\frac{d\varrho}{\pi}.$$

Suppose that this relation between L, E, A can be inverted with respect to E, for E, A in some open set V, in the form $E = \varepsilon(L, A)$ so that

$$L \equiv \lambda(\varepsilon(L, A), A)$$

with ε of class C^∞. Show that if $E = \varepsilon(L, A)$,

$$\frac{2\pi}{T_1(E, A)} = \frac{\partial \varepsilon}{\partial L}(L, A), \qquad \frac{2\pi}{T_2(E, A)} = \frac{\partial \varepsilon}{\partial A}(L, A).$$

(*Hint:* Note that $1 = \frac{\partial \lambda}{\partial E} \cdot \frac{\partial \varepsilon}{\partial L}$ and then use Eqs. (4.9.18) and (4.9.19) remarking that the derivatives with respect to the integration extremes vanish as, by the definition of ϱ_-, ϱ_+, the integrand vanishes at the extremes.)

2. In the context of Problem 1 the Hamiltonian corresponding to Eq. (4.9.8) is, ($m = 1$):

$$H(p_\varrho, p_\theta, \varrho, \theta) = \frac{1}{2}(p_\varrho^2 + \frac{p_\theta^2}{\varrho^2}) + V(\varrho)$$

and note that the function

$$\widetilde{S}(L, A, \varrho, \theta) = A\theta + \int_{\varrho_-(\varepsilon(L,A),A)} \sqrt{2(\varepsilon(L, A) - V_A(\varrho'))}\, d\varrho'$$

solves the Hamilton-Jacobi equation

$$H(\frac{\partial \widetilde{S}}{\partial \varrho}, \frac{\partial \widetilde{S}}{\partial \theta}, \varrho, \theta) = \varepsilon(L, A)$$

From this fact, infer that S generates a change of coordinates (completely canonical)

$$(p_\varrho, p_\theta, \varrho, \theta) \longleftrightarrow (L, A, \ell, g)$$

where ℓ, g are angular variables defined in Eqs. (4.9.31) and (4.9.32) in terms of the data:

$$\ell = \frac{2\pi}{T_1(\varepsilon(L, A), A)} t_0, \qquad g = \theta - S(t_0, \varepsilon(L, A), A),$$

and the Hamiltonian in the new variables is simply $H = \varepsilon(L, A)$.

3. In the context of Problem 2, define

$$\widehat{S}(E, A, \varrho, \theta) = A\theta + \int_{\varrho_-(E,A)}^{\varrho} \sqrt{(2(E - V_A(\varrho')))}\, d\varrho'.$$

Check that this is a two-parameters local solution to the Hamiltonian-Jacobi equation

$$H(\frac{\partial \widehat{S}}{\partial \varrho}, \frac{\partial \widehat{S}}{\partial \theta}, \varrho, \theta) = E$$

(the parameters being E and A) and $\widehat{S}$ generates a completely canonical transformation $(p_\varrho, p_\theta, \varrho, \theta) \longleftrightarrow (E, A, \tau, \alpha)$ in which α is a constant angle and τ varies linearly over time. Show that these new coordinates cannot be extended to a well-defined system of coordinates in the vicinity of a full trajectory of the motion if this trajectory corresponds to a quasi-periodic motion with two periods having irrational ratio. Note that this is not the case for the other coordinate transformation of Problem 2.

4. In the context of Problem 3, the change of coordinates introduced there can be extended to a well-defined system of coordinates in the vicinity of a full trajectory for which $\varrho_+(E, A) = +\infty$ (i.e., in the vicinity of an unbounded trajectory) if $\limsup_{\varrho \to +\infty} V_A(\varrho) < E$. In this case, the pair of variables (E, τ) are called "energy-time" coordinates. Why?

5. Solve Problems 1 and 2 for arbitrary $m > 0$.

4.10 Kepler's Marvelous Laws

Leva dunque, lettore, a l'alte ruote
Meco la vista, dritto a quella parte
Dove l'un moto e l'altro si percuote;
E lí comincia a vagheggiar ne l'arte
Di quel maestro che dentro a sé l'ama
Tanto che mai da lei l'occhio non parte,
Vedi come da indi si dirama
L' oblico cerchio che i pianeti porta
Per sodisfare al mondo che li chiama.[8]

The main result of §4.9, expressed by Proposition 22, is that the motion of a point mass in a central force field under some hypotheses on the initial data is a quasi-periodic motion with two periods given by Eqs. (4.9.18) and (4.9.19) depending upon the energy E and the areal velocity $\frac{1}{2}A$.

By contemplating Eqs. (4.9.18) and (4.9.19), it is easy to convince oneself that in general $T_1(E, A)$ and $T_2(E, A)$ are "independent". Hence, unless

$$\frac{T_1(E, A)}{T_2(E, A)} = \text{rational number} \tag{4.10.1}$$

which is "exceptional" when E and A vary, the motion is actually quasi periodic and not periodic.

Note, however, that the set of the space points where Eq. (4.10.1) holds will generally be dense in the region W where the motion is integrable. As an exercise, the reader may show the truth of this statement near a point of W where the E and A values are such that the Jacobian determinant of the map $(E, A) \longleftrightarrow (T_1(E, A), T_2(E, A))$ does not vanish.

However, there are two exceptional and marvelous cases.

The first, already implicitly studied in §4.1, is the harmonic oscillator bound to O by a force with potential

$$V(\varrho) = \frac{k}{2}\varrho^2 \tag{4.10.2}$$

leading to

[8] In basic English:
Look up now, reader, to the high wheels
together with me, straight there
where several motions hit each other.
And there begin to wonder about the art
Of that master who inside himself moves them with his love
so much that he never drops his eyes away.
Look up how the oblique circle bearing the planets develops there
to satisfy the world that calls them.
(Dante, Paradiso, Canto X)

$$2T_1 \equiv T_2 = 2\pi\sqrt{\frac{k}{m}} \equiv T \tag{4.10.3}$$

and the orbits are ellipses centered in O. Equation (4.10.3) could be proved by computing the integrals of Eqs. (4.9.18) and (4.9.19) (which is a long but straightforward calculation). However, the reader should try to find a simple argument leading to Eq. (4.10.3) without any explicit calculations beyond the ones already done in §4.9.

The other case corresponds to

$$V(\varrho) = -m\frac{g}{\varrho} \tag{4.10.4}$$

This is the case of the so-called "Newtonian two-body problem" or "Kepler's problem". If $E < 0$, the motion is periodic and $T_1 = T_2$, although T_1 and T_2 now actually depend on A and E, and the orbits are ellipses with focus in O.

We treat this problem in some detail by proving the following proposition.

23 Proposition. *The motions with energy $E < 0$ and areal velocity $\frac{1}{2}|A| \neq 0$ are periodic and the integrals of Eqs. (4.9.18) and (4.9.19) coincide, $\forall\, E < 0, \forall\, A \neq 0$. Furthermore:*
(i) the trajectories $t \to \varrho(t)$, $t \in \mathcal{R}_+$, are ellipses with focus in O;
(ii) such ellipses are run with constant areal velocity $\frac{A}{2}$;
(iii) the ratio between the square of the revolution period T and the cube of the length of the ellipse major axis is a constant solely depending on g.

Finally, if ϱ_+ and ϱ_- are the focal distances of the ellipse on which a given motion takes place:

$$\varrho_+ + \varrho_- = \frac{mg}{-E}, \tag{4.10.5}$$

$$\varrho_+\varrho_- = \frac{mA^2}{-2E}, \tag{4.10.6}$$

$$T = \frac{\pi}{\sqrt{2g}}(\varrho_+ + \varrho_-)^{\frac{3}{2}}. \tag{4.10.7}$$

Observations.
(1) (i), (ii), and (iii) are Kepler's laws. Starting from them, Newton realized that if one wanted to describe the motion of a planet by a second-order differential equation $m\ddot{\boldsymbol{\varrho}} = \mathbf{F}(\boldsymbol{\varrho})$, the only possibility was that $\mathbf{F}(\boldsymbol{\varrho}) = -\frac{mg}{\varrho^2}\frac{\boldsymbol{\varrho}}{\varrho}$. This led him to assume, by symmetry, $g = kM$, $M =$ mass of the Sun, i.e., $V(\varrho) = -k\frac{mM}{\varrho}$ which is the universal law of gravitation.
Of course, he also assumed that (i), (ii), and (iii) would describe the motion laws of an arbitrary body revolving around the Sun, whatever its initial position and speed.
Newton's argument is interesting and different in spirit from the one based on

the analytic theory of differential equations. It is based on some beautiful geometric considerations relying on the theory of conic sections: it can be found in the first book of the *Principia*, [37].
(2) One could also easily study the $E > 0$ or $E = 0$ motions: they are not periodic motions and the trajectories become a hyperbola wing or a parabola, respectively. This is a simple exercise along the lines of the upcoming proof and it will be left to the reader.
(3) The heavenly bodies have finite extension. Hence, if a satellite revolves circularly around a primary body (planet or Sun), turning always the same face to it, a situation in apparent contradiction to Kepler's laws is produced. In fact, if the satellite is thought of as decomposed into small point masses, the points on one face rotate on an orbit with radius smaller than the orbit of the points of the opposite face. Hence, if one could neglect the mutual interactions between the points of the body, they would have to have a different rotation period around the main body (by the Kepler's third law) and the satellite would disintegrate over time. This means that if the above catastrophic event does not occur, the body must be subject to some internal stresses ("tidal stresses") which cannot be stronger "than the body's material resistance" (otherwise, the satellite could not exist). So Kepler's laws and the gravitation law provide a mechanism for explaining Saturn rings and why, in general, satellites stay quite far away from a planet (see problems at the end of this Section).

Proof. With the notation of §4.9, let $V_A(\varrho) = m(\frac{A^2}{2\varrho^2} - \frac{g}{\varrho})$ and the angular momentum and energy conservation laws lead to

$$\varrho^2\dot\theta = A, \qquad \frac{1}{2}(\dot\varrho^2 + \frac{A^2}{2\varrho^2}) - \frac{g}{\varrho} = \frac{E}{m} \tag{4.10.8}$$

or

$$\dot\varrho^2 = \frac{2}{m}(E - V_A(\varrho)), \tag{4.10.9}$$

The graph of $V_A$is illustrated in Fig. 4.8.

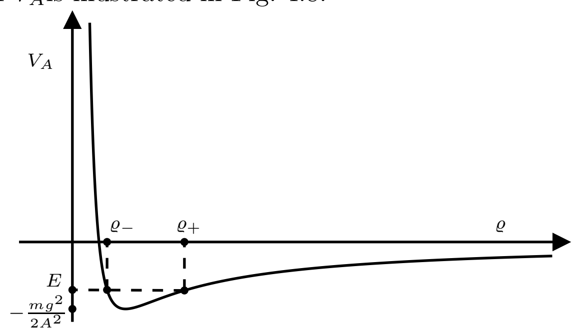


Figure 4.8: Gravitational potential in presence of the "centrifugal barrier" $\frac{mA^2}{2\varrho^2}$.

Hence, if $E < 0$, $E > -\frac{mg^2}{2A^2}$ the roots of the equation $V_A(\varrho) = E$ are $\varrho_-(E,A), \varrho_+(E,A)$ and they can be explicitly found by solving a second-degree equation in the unknown $\frac{1}{\varrho}$. Factorizing the polynomial in $\frac{1}{\varrho}$ given by $\frac{E}{m} - \frac{V_A(\varrho)}{m}$ in terms of its roots $\frac{1}{\varrho_\pm}$:

$$\frac{E}{m} - \frac{A^2}{2\varrho^2} + \frac{g}{\varrho} = \frac{A^2}{2}\big(\frac{1}{\varrho_-} - \frac{1}{\varrho}\big)\big(\frac{1}{\varrho} - \frac{1}{\varrho_+}\big) \tag{4.10.10}$$

The radii $\varrho_- < \varrho_+$ are, as we shall shortly see, the focal distances of the ellipse on which the motion develops. They obviously verify Eqs. (4.10.5) and (4.10.6) because $\varrho_+^{-1} + \varrho_-^{-1} = \frac{2g}{A^2}$, $\varrho_+\varrho_- = -\frac{2E}{mA^2}$.

By Eq. (4.10.10) to rewrite Eqs. (4.10.8) and (4.10.9) as

$$\dot\varrho = \pm A\sqrt{(\frac{1}{\varrho_-} - \frac{1}{\varrho})(\frac{1}{\varrho} - \frac{1}{\varrho_+})}, \tag{4.10.11}$$

$$\dot\theta = \frac{A}{\varrho^2}. \tag{4.10.12}$$

and suppose that for $t = 0$, it is $\varrho(0) = \varrho_-$, $\theta(0) = \pi$. Since the motion of p is periodic, being a solution to Eq. (4.9.14), and oscillates between ϱ_- and ϱ_+, this hypothesis does not affect the generality.

Then in Eq. (4.10.11) the + sign holds for $t \in [0, \frac{T}{2}]$ if T is the period of the ϱ-motion [Eq. (4.9.18)]:

$$T(E,A) = 2\int_{\varrho_-(E,A)}^{\varrho_+(E,A)} \frac{d\varrho}{\sqrt{\frac{2}{m}(E - V_A(\varrho))}}. \tag{4.10.13}$$

Hence, for $t \in [0, \frac{T}{2}]$, Eq. (4.10.12) implies that θ is a strictly increasing function (as $A > 0$ by assumption) of t: thus ϱ can be regarded as a function of θ instead of t so that Eq. (4.10.11) divided by Eq. (4.10.12) yields

$$\frac{d\varrho}{d\theta} = \varrho^2\sqrt{(\frac{1}{\varrho_-} - \frac{1}{\varrho})(\frac{1}{\varrho} - \frac{1}{\varrho_+})}. \tag{4.10.14}$$

For $\varrho_- < \varrho < \varrho_+$ this implies

$$\theta - \pi = \int_{\varrho_-}^{\varrho} \frac{d\varrho'}{\varrho'^2\sqrt{(\frac{1}{\varrho_-} - \frac{1}{\varrho'})(\frac{1}{\varrho'} - \frac{1}{\varrho_+})}}, \tag{4.10.15}$$

which is an elementary integral. Changing the variable as $y = \varrho^{-1}$, after some algebra, one finds

$$\frac{1}{\varrho} = \frac{1}{2}(\frac{1}{\varrho_+} + \frac{1}{\varrho_-}) + \frac{1}{2}(\frac{1}{\varrho_-} - \frac{1}{\varrho_+})\cos(\theta - \pi), \tag{4.10.16}$$

showing that when θ reaches 2π, ϱ reaches ϱ_+.

The study of the trajectory for $t \in [T/2, T]$ proceeds likewise, changing the choice of sign in Eq. (4.10.11), and one finds that the trajectory still verifies Eq. (4.10.16), and at time T when ϱ takes the value ϱ_-, the angle θ takes the value 3π. This means that after time T has elapsed, not only ϱ but also θ take on the initial value (of course θ has to be measured $\mod 2\pi$). Hence, the trajectory is closed because $\dot\theta$ and $\dot\varrho$ also take on again the initial values, by Eqs. (4.10.11) and (4.10.12) (i.e., $\dot\varrho = 0, \dot\theta = \frac{A}{\varrho_-^2}$) and because of the autonomy of the equations of motion.

Equation (4.10.16), well known from elementary geometry, is the polar coordinates equation of an ellipse with focus at the origin, focal distances ϱ_- and ϱ_+, and major axis along the x-axis (and "perihelion" on the negative x-axis).

To compute the period of the motion, it suffices to calculate the integral of Eq. (4.10.13), elementary after the substitution $y = \varrho^{-1}$. However, this calculation can be avoided by recalling that the ellipse is run with constant areal velocity $\frac{A}{2}$ and, hence, T can be obtained by dividing the area of the ellipse of Eq. (4.10.16) by $\frac{A}{2}$. This area is

$$\pi \frac{\varrho_+ + \varrho_-}{2} \sqrt{\varrho_+ \varrho_-} \tag{4.10.17}$$

because the semi-axes of an ellipse with focal distances ϱ_+ and ϱ_- are $\frac{\varrho_+ + \varrho_-}{2}$ and $\sqrt{\varrho_+ \varrho_-}$. Hence,

$$T = \pi \frac{\varrho_+ + \varrho_-}{2} \sqrt{\varrho_+ \varrho_-} \frac{2}{A} = \frac{\pi}{\sqrt{2g}} (\varrho_+ + \varrho_-)^{\frac{3}{2}} \tag{4.10.18}$$

by Eqs. (4.10.5) and (4.10.6). mbe

4.10.1 Exercises and Problems

Use the tables in Appendix P for the numerical values, when necessary. Problems 1 through 9 are inspired from [6].

1. Let T' be a heavenly body identical to the Earth. Could a satellite T'' identical to the Earth (i.e., a twin) be eternally eclipsed by T' while they revolve around the Sun S on a circular orbit in a one-year period? Compute the $T'T''$ distance as well as the ST' distance, comparing the percentage difference between ST' and the actual average distance between the Sun and the Earth.

2. Could a point mass M have two homogeneous rigid gravitational satellites with radius $\frac{\delta}{2}$ and mass μ whose surfaces touch at a point at distance ϱ from M? Find the necessary relations among ϱ, δ, μ, M assuming $\delta \ll \varrho$ and to first order in $\frac{\delta}{\varrho}$. Compute the force τ ("disruptive force") due to the spheres contact. (Answer: $\delta < \varrho \sqrt[3]{\frac{2}{3}\frac{\mu}{M}}$, $\tau = \frac{k\mu^2}{\delta^2}(1 - \frac{3}{2}\frac{\mu}{M}(\frac{\delta}{\varrho})^3)$, k being the gravitational constant. Suppose that the force τ cannot be negative, i.e., that the two bodies can only "push" each other.)

3. Same as Problem 2, but assuming that the body with mass M is a homogeneous sphere with radius R and that both the planet and the satellites have the same density σ: $M =$

$\frac{4\pi}{3}\sigma R^3$, $\mu = \frac{4\pi}{3}\sigma(\frac{\delta}{2})^3$. Show that to first order in $\frac{\delta}{\varrho}$ there is no condition on δ but only a condition on the ratio between R and ϱ. (*Answer:* $\tau > 0 \longleftrightarrow 1 - \frac{3}{2}8 \cdot (\frac{R}{\varrho})^3 > 0 \longleftrightarrow \varrho > 2.29R$.)

4. Use Problem 3 to show that a heuristic estimate for the minimum distance of a planet to the Sun center is $\sim 2.29R$ if R is the Sun radius. Compute ϱ in km and compare it to the orbital radius of Mercury, schematizing the Sun as a sphere with radius equal to its optically apparent radius.

5. Same as Problem 4 to estimate at what distance from the Earth can one find the closest satellite with the same density ($\sim 2.29 \times 6.3 \times 10^3\, km$). Why can the artificial satellites gravitate much closer? (See Exercise 6.)

6. Assume that a satellite to a planet is made of rock with density σ, cohesion force per unit surface γ, and with diameter δ. Using Problem 3, find a heuristic estimate of how large must δ be in order that the satellite cannot gravitate at distance ϱ from the planet (supposed to have the same density) if ϱ is in the forbidden band ($\varrho < 2.29R$). (*Hint:* Compute the tidal force τ of Exercise 3 and compare it with the cohesion force $\pi(\frac{\delta}{2})^2\gamma$: if $[k\frac{4\pi}{3}(\frac{\delta}{2})^3](\frac{3}{2}8(\frac{R}{\varrho})^3 - 1) > \varrho(\frac{\delta}{2})^2\gamma$ the tidal force prevails over the cohesion force and the body breaks up.)

7. Let $\sigma = 5.5\, g/cm^3$, $\gamma = 100\, kg_w/cm^2$, $k = 6.67 \times 10^{-8}\, cm^3/g \cdot sec^3$, $\varrho = 7.0 \times 10^3\, km$, $R = 6.33 \times 10^3\, km$. How big should a rocky satellite be in order to apply the instability argument of Problem 4? Same for $\varrho = 2R$. ($1\, kg_w$ = weight of a mass of $1\, kg$ at the Earth surface.)

8. At what distance from Saturn can one find its closest satellite? Compare it with the distance of Mimas.

9. Assuming that Saturn rings consist of rocky satellites with a cohesion modulus γ like that of Exercise 7 and a density equal to that of Saturn ($3.g/cm^3$), heuristically estimate how big can the rings stones be as a function of the radius r of the ring. Compare their maximum diameter with the observed width of the rings ($\sim 20\, km$).

10. Solve explicitly Problems 1, 2, and 4 in §4.9 in the case of Kepler's problem, explicitly computing L and $\varepsilon(L, A)$. (*Answer:* $\varepsilon(L, A) = -\frac{k^2 m^3}{2(L+mA)^2}$ if $V(\varrho) = -\frac{km}{\varrho}$.)

11. Given a Kepler motion in $\mathcal{R}^3$ with energy E, set $a = \frac{\varrho_+ + \varrho_-}{2}$, $e = \frac{\varrho_+ - \varrho_-}{\varrho_+ + \varrho_-}$ = (eccentricity of the ellipse with focal distances ϱ_+ and ϱ_-), and set

$$L = m\sqrt{k}\sqrt{a} \equiv \frac{m^{\frac{3}{2}} k}{\sqrt{-2E}}, \qquad G = L\sqrt{1 - e^2} \equiv mA.$$

Applying Problems 1, 2, and 5 of §4.9, consider the canonical transformation $(p_\varrho, p_\theta, \varrho, \theta) \longleftrightarrow (L, G, \ell, g)$ associated with the generating function

$$\widetilde{S}(L, G, \varrho, \theta) = \theta G + \int_{\varrho_-}^{\varrho} \sqrt{2m(\varepsilon(L) - V_{G/m}(\varrho))}\, d\varrho,$$

where $\varepsilon(L) = -\frac{k^2 m^3}{2L^2} = E$ and ϱ_+ and ϱ_- depend on L, G (being equal to $\varrho_+(E, A)$, and $\varrho_-(E., A)$) i.e., consider the map I generated by

$$p_\varrho = \frac{\partial \widetilde{S}}{\partial \varrho}, \quad \ell = \frac{\partial \widetilde{S}}{\partial L}, \quad p_\theta = \frac{\partial \widetilde{S}}{\partial \theta}, \quad g = \frac{\partial \widetilde{S}}{\partial G}.$$

Applying Problems 1, 2, and 5 of §4.9 and Problem 10 above, show that I can be extended to the entire set of initial data such that $G > 0, E_G \equiv -\frac{m^3 k^2}{2G^2} < E < 0$ and that the image

of this set of data via I has the form $V\times\mathcal{T}^2$, where $V=\{(L,G)\,|\,G>0, E_G<-\frac{m^3k^2}{2G^2}<0\}\equiv\{(L,G)\,|\,G>0,\,L>0\}$, and check that ℓ, g are "angles", $(\ell,g)\in\mathcal{T}^2$.

12. Show that the physical interpretation of the angle g canonically conjugated to G in Problem 11 is that of the longitude of the major semiaxis of the ellipse, while the angle ℓ conjugated to L, "average anomaly", is $\ell=\frac{2\pi}{T}t$, where t is the time necessary to reach the initial point of the orbit starting, say, at time zero from the "perihelion", i.e. from the extreme point on the major axis closest to the center of force.

13. Consider a point attracted to the origin by a gravitational force. Suppose that its energy is negative so that it moves on an ellipse and let (L,G,ℓ,g) be its Keplerian coordinates (see Problems 11 and 12). Let $(p_\varrho,p_\theta,\varrho,\theta)$ be the corresponding "natural canonical coordinates" (see Problem 2, §4.9) and let β be the polar angle formed by the position vector with the major semiaxis of the ellipse on which the motion develops following the initial data $(p_\varrho,p_\theta,\varrho,\theta)$. Call a, b and e the major semiaxis, the minor semiaxis, and the eccentricity of the ellipse, respectively, and write its equation as

$$\varrho=\frac{p}{1-e\cos\theta},\qquad p\stackrel{def}{=}\frac{b^2}{a}=\frac{2\varrho_+\varrho_-}{\varrho_++\varrho_-}$$

[see Eq. (4.10.16)] and define ξ, the "eccentric anomaly", as

$$\varrho=a\,(1+e\,\cos\xi).$$

Find relations expressing ϱ,θ,β in terms of the Keplerian variables (L,G,ℓ,g). Show that

$$\begin{aligned}
\ell=&\sqrt{1-e^2}^3\int_0^\beta\frac{d\beta'}{(1-e\cos\beta')^2}=\beta+2e\sin\beta+\frac{3}{4}\sin 2\beta+\ldots,\\
\beta=&\ell-2e\sin\ell+\frac{5}{4}e^2\sin 2\ell+\ldots,\\
\ell=&\xi+e\sin\xi,\\
\xi=&\ell-e\sin\ell+\frac{e^2}{2}\sin 2\ell+\ldots,\\
\theta=&g+\beta,\\
\varrho=&p\,(1-e\cos\theta)^{-1}=a\,(1+e\cos\xi).
\end{aligned}$$

For a more detailed theory of the equation $\ell=\xi+e\sin\xi$ see problem 6, p.486 where the radius of convergence of the inverse function, the "Laplace limit" is discussed. (*Hint:* Use Eq. (4.10.12) and $\ell=\frac{2\pi}{T}t$ to find that $\frac{d\beta}{d\ell}=\frac{(1-e\cos\beta)^2}{\sqrt{1-e^2}^3}$ (noting that β is the analogue of the angle θ of §4.10) and having used all the relations between ϱ_+,ϱ_-,A,T in Eqs. (4.10.5)-(4.10.7). Use Eq. (4.10.11) to see analogously that

$$\frac{d\varrho}{d\ell}=\frac{a}{\varrho}\sqrt{a^2e^2-(\varrho-a)^2}.$$

Then integrate the first equation to express ℓ in terms of β and the second equation to express ℓ in terms of ξ after changing variables as $\varrho=a\,(1+e\cos\xi)$.
To prove the expansion for ℓ as an eccentricity series, consider the integral expression of ℓ in terms of β found above and expand the function $\frac{\sqrt{1-e^2}^3}{(1-e\cos\beta')^2}$ in powers of e before integrating and, then, integrate term by term.)

14. Using Problem 13, express the Cartesian coordinates of the position in terms of (L,G,ℓ,g), proving that

$$x = a\,(1 + e\cos\xi)\,\cos(g+\beta), \qquad y = a\,(1 + e\cos\xi)\,\sin(g+\beta),$$

or

$$x = \frac{p}{1 - e\cos\beta}\cos(g+\beta), \qquad y = \frac{p}{1 - e\cos\beta}\sin(g+\beta),$$

where ξ, β have to be expressed in terms of (L, G, ℓ, g) via the formulae of Problem 13.

15. Using Problems 13 and 14, show that the Cartesian coordinates can be expressed correctly in terms of (L, G, ℓ, g) up to second order in the eccentricity e as

$$\begin{aligned} x =& a[\cos(g+\ell) + eA_x(g,\ell) + e^2 B_x(g,\ell)] + O(e^3), \\ y =& a[\sin(g+\ell) + eA_y(g,\ell) + e^2 B_y(g,\ell)] + O(e^3), \end{aligned}$$

where

$$\begin{aligned} A_x =& \cos g + \sin\ell\sin(g+\ell), \qquad A_y = \sin g - \sin\ell\cos(g+\ell), \\ B_x =& \sin g\sin\ell - \frac{3}{4}\sin(g+\ell)\sin 2\ell, B_y = -\cos g\sin\ell + \frac{3}{4}\cos(g+\ell)\sin 2\ell, \end{aligned}$$

which, calling ε the eccentricity to avoid confusion with $e = 2.71\ldots$, can also be written in complex form

$$x + iy = a\,e^{i(g+\ell)}[1 + (\varepsilon(e^{-i\ell} - i\sin\ell) + \varepsilon^2(-i(\sin\ell)e^{-i\ell} + \frac{3}{4}i\sin 2\ell)].$$

16. Consider the problem analogous to Problems 10 and 11 in the case of the Kepler motion in $\mathcal{R}^3$ and look for a completely canonical transformation between the natural polar coordinates $(p_\varrho, p_\varphi, p_\theta, \varrho, \varphi, \theta)$ in terms of which the Hamiltonian is

$$\frac{1}{2m}(p_\varrho^2 + \frac{p_\varphi^2}{(\varrho\sin\theta)^2} + \frac{p_\theta^2}{\varrho^2}) - \frac{km}{\varrho},$$

(here ϱ =radial distance, φ =longitude and θ =latitude) corresponding to the Lagrangian

$$\frac{m}{2}(\dot\varrho^2 + \varrho^2(\sin\theta)^2\dot\varphi^2 + \varrho^2\dot\theta^2) + k\frac{m}{\varrho}$$

and the coordinates $(L, G, \Theta, \ell, g, \tau)$, where L, G, Θ are defined in terms of the energy E, the areal velocity A and of the orbit inclination i with respect to the z-axis by

$$L\frac{m^{\frac{3}{2}}k}{\sqrt{-2E}}, \quad G = mA, \quad \Theta = G\cos i$$

and ℓ, g, τ are their canonically conjugated variables which will turn out to be $\ell =$ (average anomaly in the ellipse plane), $g =$ (longitude of the major semiaxis of the ellipse in its plane measured, say, from the nodal line, i.e., from the line of intersection of the ellipse plane and the $(\mathbf{i}, \mathbf{j})$-plane of the inertial reference system $(0; \mathbf{i}, \mathbf{j}, \mathbf{k})$ to which the motion is referred), $\tau =$ (longitude of the nodal line of the ellipse plane in the plane $(\mathbf{i}, \mathbf{j})$ measured, say, from $\mathbf{i}$, "angle or ascension").
Show that, if $\varepsilon(L) = \frac{m^3k^2}{-2L^2}$, the above transformation is completely canonical and is generated by the solution of the Hamilton-Jacobi equation

$$\frac{1}{2m}\left((\frac{\partial S}{\partial\varrho})^2 + \frac{1}{\varrho^2\sin^2\theta}(\frac{\partial S}{\partial\varphi})^2 + \frac{1}{\varrho^2}(\frac{\partial S}{\partial\theta})^2\right) - \frac{km}{\varrho} = \varepsilon(L),$$

parameterized by L, G, Θ and having the form (solution with "separation of variables")

$$S(\varrho,\theta,\varphi;L,G,\Theta) = -(\frac{\pi}{2}-\varphi)\,\Theta + \widetilde{\sigma}_{G,\Theta}(\theta) + \widetilde{\sigma}_{G,L}(\varrho) - \frac{\pi}{2}G$$

with

$$\Big(\frac{d\sigma_{G,\Theta}}{d\theta}\Big)^2 = G^2 - \frac{\Theta^2}{\sin^2\theta},$$
$$\Big(\frac{d\sigma_{G,\Theta}}{d\varrho}\Big)^2 = 2m\Big(\varepsilon(L) - V_{G/m}(\varrho)\Big) = 2m\Big(-\frac{m^3k^2}{2L} - \frac{G^2}{2m^2\varrho^2} + k\frac{m}{\varrho}\Big).$$

(*Hint:* The latitude θ varies between $\theta_- = \frac{\pi}{2} - i$ and $\theta_+ = \frac{\pi}{2} + i$, assuming to have chosen the axis normal to the ellipse plane oriented so that $i < \frac{\pi}{2}$. The ϱ variable varies between ϱ_- and ϱ_-. Then $\theta_-, \theta_+, \varrho_-, \varrho_+$ can be computed from L, G, Θ. Write

$$\widetilde{\sigma}_{G,\Theta}(\theta) = \int_{\theta_-}^{\theta} \sqrt{G^2 - \frac{\Theta^2}{\sin^2\theta'}}d\theta', \qquad \widetilde{\sigma}_{G,L}(\varrho) = \int_{\varrho_-}^{\varrho} \sqrt{2m(\varepsilon(L) - V_{G/m}(\varrho'))}d\varrho'$$

and note that the variables τ, g, ℓ are defined by

$$\tau = -\frac{\pi}{2} + \varphi + \frac{\partial\widetilde{\sigma}_{G,\Theta}}{\partial\Theta}(\theta), \qquad g = \frac{\partial\widetilde{\sigma}_{G,\Theta}}{\partial G}(\theta) + \frac{\partial\widetilde{\sigma}_{G,L}}{\partial G}(\varrho) - \frac{\pi}{2}, \qquad \ell = \frac{\partial\widetilde{\sigma}_{G,L}}{\partial L}(\varrho).$$

In the new variables, the Hamiltonian becomes $\varepsilon(L)$ so that the Keplerian evolution is, in these variables, τ = constant, g = constant. So we can compute τ and g by choosing special phase-space points on the orbit. To find the meaning of τ, consider the time when the point occupies the "highest position", $\theta = \theta_-$. Show that $\frac{\partial\widetilde{\sigma}}{\partial\Theta}(\theta_-) = 0$, noting that the argument of the integral for $\widetilde{\sigma}_{G,\Theta}$ vanishes for $\theta = \theta_\pm$. Hence, $\tau = -\frac{\pi}{2} + \varphi$ when $\theta = \theta_-$. Geometrically, this means that τ is the angle formed with the x-axis by the line in the xy plane orthogonal to the projection on the xy plane of the normal to the orbital plane, i.e., the nodal line.
Similarly, to find the meaning of g, consider the time when $\varrho = \varrho_-$ (i.e., the point is at the perihelion). Now, $\frac{\partial\widetilde{\sigma}_{G,L}}{\partial G}(\varrho_-) = 0$ and

$$g = \frac{\Theta^2}{G^2}\frac{1}{\sin^2\theta} = 1 - \frac{\cos^2 i}{\sin^2\theta}$$

where θ_0 is the polar angle corresponding to the perihelion position. This relation can be interpreted as saying that $g - \frac{\pi}{2}$ is the time necessary for a point moving according to the equation

$$\dot{\theta}^2 = 1 - \frac{\Theta^2}{G^2}\frac{1}{\sin^2\theta}$$

to go from θ_- to θ_0. On the other hand, it is easy to see that the above equation also describes the θ variation over time in a circular uniform motion on the unit circle in the plane of the ellipse (inclined by i) with unit speed. So $g - \frac{\pi}{2}$ is the angle between the major semiaxis of the ellipse and the intersection between the ellipse plane and the plane containing the normal to the ellipse and the z-axis ("azimuthal plane" of the normal), since the angle between the latter line and the nodal line is $\frac{\pi}{2}$, it follows that g has the desired interpretation. The angle ℓ has the same expression found for the planar case (see problem 11). Hence it has the same interpretation of average anomaly).

17. Express the Cartesian coordinates of the position corresponding to $(L, G, \Theta, \ell, g, \tau)$ of Problem 16. (*Hint:* Use the results of Problem 15 directly.)

4.11 Integrable Systems. Solid with a Fixed Point

Consider N point masses, with masses $m_1, \ldots, m_N > 0$, subject to an ideal constraint imposing that the system be rigid and have a fixed point O. Suppose $N \geq 3$ and that the points are not aligned.

We shall describe the motions in a reference frame $(O; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \bar{\mathbf{k}})$, conventionally called "fixed", and we shall fix a "comoving" frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with axes suitably chosen.

To determine the position of the body, it will suffice to give the position of the reference frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ since, in this system of coordinates, the i-th point has constant coordinates by the rigidity constraint. We shall use the Euler angles $(\overline{\theta}, \overline{\varphi}, \overline{\psi})$ to define $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$; they are defined in §3.9, Fig. 3.3 (see Fig. 4.9):

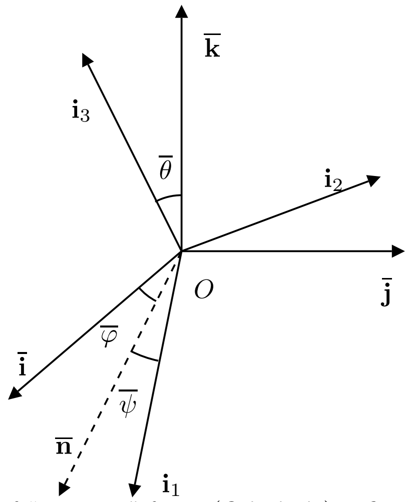


Figure 4.9. Euler angles of "comoving" frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ in fixed frame $(O; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \bar{\mathbf{k}})$.

The kinetic energy can be expressed in terms of the angular velocity $\boldsymbol{\omega}$ of $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to $(O; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \bar{\mathbf{k}})$, [see Eqs. (3.9.11) and (3.9.12)]:

$$\boldsymbol{\omega} = \dot{\overline{\theta}}\,\bar{\mathbf{n}} + \dot{\overline{\varphi}}\,\bar{\mathbf{k}} + \dot{\overline{\psi}}\,\mathbf{i}_3 \tag{4.11.1}$$

In fact, the velocity of the i-th point can simply be written as

$$\dot{\mathbf{x}}^{(i)} = \boldsymbol{\omega} \wedge (P_i - O) \tag{4.11.2}$$

(see footnote 10, p.202, last formula). Therefore

$$\begin{aligned} T =& \frac{1}{2}\sum_{i=1}^{N} m_i (\dot{\mathbf{x}}^{(i)})^2 = \frac{1}{2}\sum_{i=1}^{N} m_i (\boldsymbol{\omega} \wedge (P_i - O)) \cdot (\boldsymbol{\omega} \wedge (P_i - O)) \\ =& \frac{1}{2}\boldsymbol{\omega} \cdot \sum_{i=1}^{N} (P_i - O) \wedge (\boldsymbol{\omega} \wedge (P_i - O)) = \frac{1}{2}\boldsymbol{\omega} \cdot I\boldsymbol{\omega}, \end{aligned} \tag{4.11.3}$$

by vector calculus, see Eq. (3.9.15), where

$$I\boldsymbol{\omega} = \sum_{i=1}^{N} (P_i - O) \wedge (\boldsymbol{\omega} \wedge (P_i - O)) = \mathbf{K}_O. \tag{4.11.4}$$

The components of $I\boldsymbol{\omega}$ in the co-moving frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ have, by Eq. (4.11.4), the form

$$(I_{\boldsymbol{\omega}})_\alpha = \sum_{\beta=1}^{3} I_{\alpha\beta}\omega_\beta, \qquad \alpha = 1, 2, 3 \tag{4.11.5}$$

and from Eq. (4.11.4) it is easy to check that

$$I_{\alpha\beta} = \sum_{i=1}^{N} m_i \, [(P_i - O)^2 \delta_{\alpha\beta} - (P_i - O)_\alpha (P_i - O)_\beta], \tag{4.11.6}$$

for instance by using the identity $\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\,\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\,\mathbf{c}$. Since the components of $(P_i - O)$ in $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ are constants, by the rigidity constraint, the nine numbers of Eq. (4.11.5), actually six since $I_{\alpha\beta} \equiv I_{\beta\alpha}$ are characteristic constants of the body associated with the frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$.

At this point, it is convenient to choose the co-moving frame so that the matrix I ("inertia matrix") is as simple as possible.

Note that by rotating the $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ axes to $\mathbf{i}'_1, \mathbf{i}'_2, \mathbf{i}'_3$ the coordinates of the vectors $(P_i - 0)$ become $(P_i - O)'_\alpha$, $\alpha = 1, 2, 3$, in the new frame, related to the old coordinates by

$$(P_i - O)_\alpha = \sum_{\beta=1}^{3} R_{\alpha\beta} (P_i - O)'_\beta \tag{4.11.7}$$

and R is an orthogonal matrix $RR^T = R^T R = 1$ $(R_{\alpha\beta} = \mathbf{i}_\alpha \cdot \mathbf{i}'_\beta)$. And, vice versa, any orthogonal matrix corresponds to some frame $(O; \mathbf{i}'_1, \mathbf{i}'_2, \mathbf{i}'_3)$ so that Eq. (4.11.7) gives expresses the change of coordinates.

Therefore, the inertia matrix depends on the co-moving frame and in $(O; \mathbf{i}'_1, \mathbf{i}'_2, \mathbf{i}'_3)$ it becomes I' related to I by

$$I = RI'R^T \tag{4.11.8}$$

by Eqs. (4.11.7) and (4.11.6), in matrix notations. Then we can choose R so that I' becomes

$$I' = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{4.11.9}$$

where $0 < I_1 \le I_2 \le I_3$. Such an R exists because I is a symmetric positive definite matrix[9] and every such matrix can be "diagonalized" by an orthogonal transformation (see Appendix F).

Hence it is not restrictive to suppose, since the beginning, that the choice of the comoving frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ is such that I takes the form of Eq. (4.11.9).

With this choice of the co-moving axes, the kinetic energy and the angular momentum become [see Eqs. (4.11.3), (4.11.4), and (4.11.5)]

$$T = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2), \tag{4.11.10}$$

$$\mathbf{K}_O = I_1\,\omega_1\,\mathbf{i}_1 + I_2\,\omega_2\,\mathbf{i}_2 + I_3\,\omega_3\,\mathbf{i}_3. \tag{4.11.11}$$

To write the Lagrangian function describing the motion of the body, with O fixed and subject to no force other than that of the ideal constraints of fixed O and of rigidity, it will be enough to express the kinetic energy in terms of the Euler angles (Fig. 4.9) and of their time derivatives, through Eqs. (4.11.1) and (4.11.10). The components of $\boldsymbol{\omega}$ become explicitly

$$\omega_1 = \dot{\overline{\theta}}\cos\overline{\psi} + \dot{\overline{\varphi}}\sin\overline{\theta}\,\sin\overline{\psi} \tag{4.11.12}$$

$$\omega_2 = \dot{\overline{\theta}}\sin\overline{\psi} + \dot{\overline{\varphi}}\sin\overline{\theta}\,\cos\overline{\psi} \tag{4.11.13}$$

$$\omega_3 = \dot{\overline{\varphi}}\cos\overline{\theta} + \dot{\overline{\psi}} \tag{4.11.14}$$

by Eqs. (3.9.3) and (4.11.1). The result is not particularly illuminating in the general case and we write it only in the "gyroscope case" when, say, $I_1 = I_2 \stackrel{def}{=} I$. One finds

$$\mathcal{L} = \frac{1}{2}\,I\,(\dot{\overline{\theta}}^2 + \sin^2\overline{\theta}\,\dot{\overline{\varphi}}^2) + \frac{1}{2}\,I_3\,(\dot{\overline{\varphi}}\cos\overline{\theta} + \dot{\overline{\psi}})^2 \tag{4.11.15}$$

Before treating the general case, let us study the system described by Eq. (4.11.15), i.e., the gyroscope. In this case, the results are easier and particularly suggestive.

As is often the case, it is not convenient to write down only the Lagrange equations for Eq. (4.11.15) and discuss them. It is better to combine them with other information which can be obtained by general conservation principles (of energy and angular momentum, in the present case). Such information, although implicitly present in the Lagrange equations, is not very obvious there.

Since $\mathbf{K}_O$ is a constant of the motion, given a motion $t \to (\overline{\theta}(t), \overline{\varphi}(t), \overline{\psi}(t))$ with initial datum $(\dot{\overline{\theta}}_0, \dot{\overline{\varphi}}_0, \dot{\overline{\psi}}_0, \overline{\theta}_0, \overline{\varphi}_0, \overline{\psi}_0)$, we can suppose without affecting generality that $\mathbf{K}_O$ is parallel to some fixed axis $\mathbf{k}$:

$$\mathbf{K}_O = A\,\mathbf{k}, \qquad A > 0. \tag{4.11.16}$$

[9] Since $\frac{1}{2}\boldsymbol{\omega}\cdot I\boldsymbol{\omega} =$ (kinetic energy of the body) $\ge 0, \forall\,\boldsymbol{\omega} \in \mathcal{R}^3$, and it can vanish only if $\boldsymbol{\omega} = \mathbf{0}$ because the points are assumed to be not aligned, see Eq. (4.11.2).

(the $A = 0$ case corresponds to a motionless solid which remains such forever, of course.)

Let $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ be a reference frame with z-axis oriented as $\mathbf{k}$ and choose $\mathbf{i}$ on the intersection between the $(\mathbf{i}, \mathbf{j})$ plane and the $(\bar{\mathbf{i}}, \bar{\mathbf{j}})$. We suppose that such planes do not coincide (otherwise, we change $(\bar{\mathbf{i}}, \bar{\mathbf{j}})$).

The motion in this new fixed frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, *whose definition however depends on the initial data*, will be discussed calling (θ, φ, ψ) the Euler angles of $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to the frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$.

The components of $\mathbf{K}_O = A\,\mathbf{k}$ in the co-moving frame are expressed. see Eq. (3.9.3), in terms of the new Euler angles as

$$(\mathbf{K}_O)_3 = A\cos\theta, \quad (\mathbf{K}_O)_2 = A\,\sin\theta\sin\psi, \quad (\mathbf{K}_O)_3 = A\,\sin\theta\cos\psi. \tag{4.11.17}$$

By relations like Eqs. (4.11.12)-(4.11.14), written with the new angles, the angular momentum conservation gives the following relations:

$$A\cos\theta = I_3\omega_3 = I_3(\dot\varphi\cos\theta + \dot\psi),) \tag{4.11.18}$$

$$A\sin\theta\cos\psi = I\omega 2, \tag{4.11.19}$$

$$A\sin\theta\sin\psi = I\omega_1, \tag{4.11.20}$$

which are three differential equations for the three unknowns θ, φ, ψ and A is a constant [ω_1, ω_2 are also expressed in terms of the angles θ, φ, ψ and of their derivatives by relations like Eqs. (4.11.12) and (4.11.13)].

Instead of discussing the above equations, which, in principle, should be sufficient to determine the motion, we shall combine them with some of the Lagrangian equations associated with Eq. (4.11.15), written in the new θ, φ, ψ variables (i.e., without the overbars). The analysis is based on [28].

Since Eq. (4.11.15) does not explicitly depend upon φ, ψ, one deduces two conservation laws from Eq. (4.11.15) by writing the Lagrange equations corresponding to the variables φ, ψ:

$$\frac{d}{dt} I_3(\dot\varphi\cos\theta + \dot\psi) = 0, \tag{4.11.21}$$

corresponding to ψ and

$$\frac{d}{dt}\big(I\sin^2\theta\,\dot\varphi + I_3(\dot\varphi\cos\theta + \dot\psi)\cos\theta = 0 \tag{4.11.22}$$

corresponding to φ.

Equations (4.11.18)-(4.11.22) form a redundant system: but they easily determine the functions $(\theta(t), \varphi(t), \psi(t))$ in terms of the initial data.

In fact, Eq. (4.11.21) implies that $\dot\varphi\cos\theta + \dot\psi$ is constant as t varies; hence, Eq. (4.11.18) implies that $\cos\theta$ is constant, i.e.,

$$\theta(t) = \theta_0 \tag{4.11.23}$$

(remark that this holds only in the reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ chosen *after* the particular motion had been selected, which is very special, for instance $\dot{\theta} = 0$ and thus $\dot{\theta}_0 = 0$.)

Using Eqs. (4.11.23) and (4.11.21) in Eq. (4.11.22), we see that $\ddot{\varphi} = 0$, i.e.,

$$\varphi(t) = \varphi_0 + \dot{\varphi}_0 t \tag{4.11.24}$$

Then the constancy of $\dot{\varphi}$ and of θ and Eq. (4.11.21) imply that $\dot{\psi}$ is also a constant:

$$\psi(t) = \psi_0 + \dot{\psi}_0 t \tag{4.11.25}$$

Hence, Eqs. (4.11.23)-(4.11.25) provide a full description of the motion in the chosen coordinates (which, we stress once more, is a reference frame depending on the motion itself, having z axis parallel to the constant angular momentum). It appears that the motion expressed in the Cartesian coordinates is quasi-periodic with periods

$$T_1 = \frac{2\pi}{\dot{\varphi}_0}, \quad T_2 = \frac{2\pi}{\dot{\psi}_0}. \tag{4.11.26}$$

It follows that the motion is quasi-periodic but generally not periodic, although the set of the initial data for which $\frac{\dot{\varphi}_0}{\dot{\psi}_0}$ is rational is a dense set of data lying on periodic orbits.

By Observation (5) to Definition 10, p.288, the above system should be integrable in the sense of Definition 10, p.287, on vast regions of the data space.

Let us study the general case, assuming $0 < I_1 < I_2 < I_3$ and using a method, inspired from [28], quite different from the preceding one.

As before, given a motion, the angular momentum is a constant together with the kinetic energy. This implies

$$I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^3 = 2E = \mathit{const}, \tag{4.11.27}$$

$$I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^3 = A^2 = \mathit{const}, \tag{4.11.28}$$

giving two of the three component of $\boldsymbol{\omega}$ in terms of the third:

$$\omega_1 = \pm\sqrt{\frac{(2EI_3 - A^2) - (I_3 - I_2)I_2\omega_2^2}{I_1(I_3 - I_1)}}, \tag{4.11.29}$$

$$\omega_3 = \pm\sqrt{\frac{(A^2 - 2EI_1) - (I_2 - I_1)I_2\omega_2^2}{I_3(I_3 - I_1)}}, \tag{4.11.30}$$

To find an equation allowing the determination of ω_2 one can remark that Eq. (4.11.28) contains less information than the constancy of the angular momentum as a vector.

In fact, the angular momentum conservation means [recalling Eq. (3.9.12)]

$$\mathbf{0} = \frac{d\mathbf{K}_O}{dt} = \frac{d}{dt}(I_1\,\omega_1\,\mathbf{i}_1 + I_2\,\omega_2\,\mathbf{i}_2 + I_3\,\omega_3\,\mathbf{i}_3) \tag{4.11.31}$$

$$= I_1\,\dot\omega_1\,\mathbf{i}_1 + I_2\,\dot\omega_2\,\mathbf{i}_2 + I_3\,\dot\omega_3\,\mathbf{i}_3 + I_1\,\omega_1\,\frac{d\mathbf{i}_1}{dt} + I_2\,\omega_2\,\frac{d\mathbf{i}_2}{dt} + I_3\,\omega_3\,\frac{d\mathbf{i}_3}{dt}$$

$$= I_1\,\dot\omega_1\,\mathbf{i}_1 + I_2\,\dot\omega_2\,\mathbf{i}_2 + I_3\,\dot\omega_3\,\mathbf{i}_3 + I_1\,\omega_1\,\boldsymbol{\omega}\wedge\mathbf{i}_1 + I_2\,\omega_2\,\boldsymbol{\omega}\wedge\mathbf{i}_2 + I_3\,\omega_3\,\boldsymbol{\omega}\wedge\mathbf{i}_3$$

which, written in components on $(O;\mathbf{i}_1,\mathbf{i}_2,\mathbf{i}_3)$, is

$$I_1\dot\omega_1 = (I_2 - I_3)\omega_2\omega_3, \tag{4.11.32}$$

$$I_2\dot\omega_2 = (I_3 - I_1)\omega_3\omega_1, \tag{4.11.33}$$

$$I_3\dot\omega_3 = (I_1 - I_2)\omega_1\omega_2. \tag{4.11.34}$$

These very beautiful equations are the "Euler equations" for the motion of the solid. Equation (4.11.33) together with Eqs. (4.11.29) and (4.11.30) give the equation for ω_2:

$$\dot\omega_2 = \pm\sqrt{\frac{\{(2EI_3-A^2)-(I_3-I_2)I_2\omega_2^2\}\{(A^2-2EI_1)-(I_2-I_1)I_2\omega_2^2\}}{I_1I_2^2I_3}} \tag{4.11.35}$$

and the discussion of the choice of sign in Eq. (4.11.35) leads to the usual result: initially, $\dot\omega_2$ has some sign which is kept until it vanishes, then the sign changes until the next time $\dot\omega_2$ vanishes, etc., alternating[10] (see §2.7).

Hence, recalling §2.7, Eq. (4.11.35) tells us that ω_2 varies over time as the abscissa of a point mass with mass 2, total energy 0, moving under the action of a conservative force with potential energy:

$$V_{E,A}(x) = \frac{\{(2EI_3-A^2)-(I_3-I_2)I_2x^2\}\{(A^2-2EI_1)-(I_2-I_1)I_2x^2\}}{I_1I_2^2I_3} \tag{4.11.36}$$

Therefore, $t\to\omega_2(t)$ is a C^∞-periodic function of t oscillating between two extreme values $\alpha_+(E,A),\alpha_-(E,A)$ which are the extremes of the smaller of the two intervals $(-a_1,a_1)$, $(-a_3,a_3)$ with a_j = roots of $V_{E,A}(x)=0$, $a_j>0, j=1,3$:

$$a_1(E,A) = \sqrt{\frac{2EI_3-A^2}{I_2(I_3-I_2)}}, \qquad a_3(E,A) = \sqrt{\frac{A^2-2EI_1}{I_2(I_2-I_1)}}, \tag{4.11.37}$$

provided

[10] As in the one-dimensional conservative problems, if $\dot\omega_2$ vanishes initially the choice of sign for $t>0$ and small can be inferred from the initial value of ω_2, (see §2.7).

$$a_1(E,A) \neq a_3(E,A); \tag{4.11.38}$$

otherwise, the equation$V_{E,A} = 0$ has only two solutions, $\pm a$ and $V'_{E,A}$ vanishes there so that the motion, by the analysis of §2.7, will be aperiodic.

The period of $t \to \omega_2(t)$ is

$$T_1(E,A) = 2\int_{\alpha_-(E,A)}^{\alpha_+(E,A)} \frac{dx}{\sqrt{-V_{E,A}(x)}}, \tag{4.11.39}$$

and a better expression for $\omega_2(t)$ can be obtained by defining

$$t \to \Omega_t(t,E,A), \qquad t \in \mathcal{R}, \tag{4.11.40}$$

to be the solution of $2\ddot{\Omega} = -\frac{\partial V_{E,A}}{\partial \omega_2}(\Omega)$, hence, of Eq. (4.11.35), with initial datum

$$\Omega(0,E,A) = \alpha_-(E,A), \qquad \dot{\Omega}(0,E,A) = 0. \tag{4.11.41}$$

Then

$$\omega_2(t) = \Omega(t + t_0(\omega_2(0), \dot{\omega}_2(0)), E, A), \tag{4.11.42}$$

where $t_0(\omega_2(0), \dot{\omega}_2(0))$ is the minimum time necessary in order that the solution (4.11.40) "reaches" the datum $\dot{\omega}_2(0), \omega_2(0)$. Furthermore, for $0 \le t \le \frac{1}{2}T_1(E,A)$, it is

$$t = \int_{\alpha_-(E,A)}^{\Omega(t,E,A)} \frac{dx}{\sqrt{-V_{E,A}(x)}}, \tag{4.11.43}$$

To find the motion $t \to (\theta(t), \varphi(t), \psi(t))$, we have to go back to the equations expressing the conservation of angular momentum and its identity with $A\,\mathbf{k}$, assuming, again, to have chosen a reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ with $\mathbf{k}$ and $\mathbf{K}_O$, parallel and $\mathbf{i}$ along the node line of the planes $(\mathbf{i}, \mathbf{j})$ and $(\bar{\mathbf{i}}, \bar{\mathbf{j}})$, see Eqs. (4.11.18)-(4.11.20). Now

$$I_3\omega_3 = A\,\cos\theta, \quad I_2\omega_2 = A\,\sin\theta\,\sin\psi, \quad I_1\omega_1 = A\sin\theta\,\cos\psi \tag{4.11.44}$$

tell us that

$$\theta(t) = \arccos\frac{I\omega_3(t)}{A}, \tag{4.11.45}$$

$$\psi(t) = \mathrm{arctg}\frac{I_2\omega_2(t)}{I_3\omega_3(t)} \tag{4.11.46}$$

where the determination of the arc-tangent has to be chosen so that $t \to \psi(t)$ is continuous.

From Eqs. (4.11.12) and (4.11.13) written without overbars (i.e., for the Euler angles of $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ with respect to $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, we deduce $\dot{\varphi}$:

$$\dot{\varphi} = \frac{\omega_1 \sin\psi + \omega_2 \cos\psi}{\sin\theta} = \frac{I_1\omega_1^2 + I_2\omega_2^2}{I_1^2\omega_1^2 + I_2^2\omega_2^2} \tag{4.11.47}$$

where the second equality follows from Eqs. (4.11.19), (4.11.20), and (4.11.18) (recalling that we are supposing $\mathbf{K}_O$ parallel to $\mathbf{k}$). Let

$$\Phi(t, E, A) = A\frac{I_1\Omega_1(t, E, A)^2 + I_2\Omega_2(t, E, A)^2}{I_1^2\Omega_1(t, E, A)^2 + I_2^2\Omega_2(t, E, A)^2}, \tag{4.11.48}$$

where Ω_1, is connected with Ω as ω_1 with ω_2 in Eq. (4.11.29) (note that the sign ambiguity has no relevance here). Then Eq. (4.11.47) becomes

$$\dot{\varphi} = \Phi(t + t_0(\omega_2(0), \dot{\omega}_2(0)), E, A). \tag{4.11.49}$$

Using the periodicity with period Eq. (4.11.39) of $t \to \Phi(t, E, A)$ and calling $(\chi_n(E, A))_{n\in\mathcal{Z}}$ the Fourier coefficients of this function, it is

$$\Phi(t, E, A) = \sum_{n=-\infty}^{+\infty} \chi_n(E, A)e^{\frac{2\pi i}{T_1(E,A)}t}, \tag{4.11.50}$$

and, by integrating Eq. (4.11.49),

$$\begin{aligned}\varphi(t) =& \varphi_0 + \chi_0(E, A)t \\ &+ S(t + t_0(\omega_2(0), \dot{\omega}_2(0)), E, A) - S(t_0(\omega_2(0), \dot{\omega}_2(0)), E, A),\end{aligned} \tag{4.11.51}$$

where

$$S(t, E, A) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{+\infty} \chi_n(E, A)\frac{e^{\frac{2\pi i}{T_1(E,A)}nt}}{\frac{2\pi i n}{T_1(E,A)}} \tag{4.11.52}$$

which is a C^∞-function periodic in t with period $T_1(E, A)$.

Equations (4.11.45), (4.11.40), (4.11.42), (4.11.51), (4.11.46), (4.11.29), and (4.11.30) give a complete description of the motion under investigation.

The analogy of the above results with those of the two-body problem lead to the formulation of the following proposition.

24 Proposition. *The motion of a solid with a fixed point and inertia moments $0 < I_1 < I_2 < I_3$ is integrable in the sense of Definition 10, §4.8, p.287, in a family of regions covering the region W of the data space where $A \neq 0$, $a_3(E, A) \neq a_1(E, A)$ [see Eq. (4.11.38)], and in such cases the motion is quasi-periodic with two periods:*

$$T_1(E, A) = 2\int_{\alpha_-(E,A)}^{\alpha_+(E,A)} \frac{dx}{\sqrt{-V_{E,A}(x)}}, \tag{4.11.53}$$

$$T_2(E,A) = \frac{2\pi}{\chi_0(E,A)} = \frac{\pi}{A} \frac{T_1(E,A)}{\int_{\alpha_-(E,A)}^{\alpha_+(E,A)} \frac{dx}{\sqrt{-V_{E,A}(x)}} \left[\frac{(2E-A^2)-(I_2-I_1)I_2x^2}{I_1(2E-A^2)-I_2I_3(I_2-I_1)x^2}\right]} \tag{4.11.54}$$

and $\alpha_\pm(E,A)$ *are the two positive roots of the smallest modulus of* $V_{E,A}(x) = 0$*, with* $V_{E,A}$ *being defined in Eq. (4.11.36).*
Similar (and simpler) results hold if $I_1 = I_2 \neq I_3$*,* $I_1 \neq I_2 = I_3$*,* $I_1 = I_2 = I_3$*.*

Observations.
(1) The proof of Proposition 24 is essentially a different way of stating what has already been discussed above. The analysis of this section (as well as that on the central forces) is a classical proof. Somehow, it seems unsatisfactory because it looks like "magic", with its use of redundant equations chosen, without apparent a priori logic, to reach the goal of finding explicit expressions for the motions. However, with further thought, it appears quite simple and, in particular, no need of the theory of elliptic functions emerges (a claim referring to deeper analysis of the properties of the quadratures discussed above).
(2) However, there is a deeper critique of the above deductions. It is not at all clear that the systems are canonically integrable in the sense of Definition 11, §4.8, p.289. This becomes very serious when one tries to study by the Hamilton-Jacobi theory the perturbations provoked by small conservative forces on the above simple motions. The reader will realize this problem more clearly in §5.10-§5.12, where the theory of the Hamiltonian perturbations based on the Hamilton-Jacobi equations is developed.
In the problems to §4.9 and §4.10 we have shown, however, how to deduce for the central motions complete canonical integrability from the integrability proof. Likewise, in the problems of this section, we show how to deduce canonical integrability of the solid motion from parts of the proof of the above proposition. The derivation is simple and nice, not so much because it leads very quickly to the quadrature formulae (4.11.42), (4.11.47), (4.11.46), and (4.11.52), but mainly because it achieves the proof of canonical integrability at the same time. This integrability property had always been discussed either abstractly or quite obscurely until recently when the "Deprit canonical transformation" was introduced.

Proof. We discuss the proof in some further details because it is useful to illustrate Observation (5) to Definition 10, p.288.

Let $(O; \overline{\mathbf{i}}, \overline{\mathbf{j}}, \overline{\mathbf{k}})$ be the fixed frame and let $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ be the "adapted" fixed frame chosen, once a particular motion is given, with the $\mathbf{k}$ axis parallel to the angular momentum. Suppose that $\mathbf{i}$ is parallel to the node of the planes $(\overline{\mathbf{i}}, \overline{\mathbf{j}})$ and $(\mathbf{i}, \mathbf{j})$ (i.e., to their intersection).

To determine the initial datum in the $I_1 = I_2$ case, we use the following coordinates:

(1) the angle γ between $\overline{\mathbf{i}}$ and $\mathbf{i}$;

(2) the angle δ between $\mathbf{K}_O$ and $\overline{\mathbf{k}}$;

(3) the Euler angles φ, ψ of $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$;
(4) the angular velocity variables $\dot{\varphi}$ and $\dot{\psi}$.

From the preceding analysis, it follows that the motion of the system has three prime integrals $(\delta, \dot{\varphi}, \dot{\psi})$ and, given them, it is described by the points $(\gamma.\varphi, \psi) \in \mathcal{T}^3$, and the time evolution on $\mathcal{T}^3$ is described by quasi-periodic flow with pulsations

$$\sigma_1(\delta, \dot{\varphi}, \dot{\psi}) = 0, \quad \sigma_2(\delta, \dot{\varphi}, \dot{\psi}) = \dot{\varphi}, \quad \sigma_3(\delta, \dot{\varphi}, \dot{\psi}) = \dot{\psi}, \tag{4.11.55}$$

having denoted them with σ instead of ω to avoid confusion with the above angular velocity components.

The integrating map is thus $I(\dot{\overline{\theta}}, \dot{\overline{\varphi}}, \dot{\overline{\psi}}, \overline{\theta}, \overline{\varphi}, \overline{\pi}) \longleftrightarrow (\delta, \dot{\varphi}, \dot{\psi}, \gamma, \varphi, \psi)$. It should still be checked that this map is C^∞ nonsingular and invertible on a suitable family of neighborhoods W' which, as one uses the arbitrariness of the choice of $(O; \bar{\mathbf{i}}, \bar{\mathbf{j}}, \overline{\mathbf{k}})$, cover W. We do not enter into this analysis.

In the general case $(I_1 < I_2 < I_3)$, we replace the variables $(\delta, \dot{\varphi}, \dot{\psi})$ which, with the exception of δ, are no longer conserved with the variables (δ, E, A), and we also replace the angles which no longer rotate uniformly with the exception of γ which is constant, with where $(\gamma, \widetilde{\varphi}, \widetilde{\psi})$ where

$$\widetilde{\psi} = \frac{2\pi}{T_1(E, A)} t_0(\omega_2(0), \dot{\omega}_2(0)), \qquad \widetilde{\varphi} = \varphi - S(t_0(\omega_2(0), \dot{\omega}_2(0)), E, A) \tag{4.11.56}$$

[see Eqs. (4.11.51), (4.11.52), and (4.11.42)].

By the discussion preceding Proposition 24, it appears that $\gamma, \widetilde{\varphi}, \widetilde{\psi}$ are angles rotating with pulsations

$$\sigma_1(A, E, \delta) = 0, \quad \sigma_2(A, E, \delta) = \tfrac{2\pi}{T_1(E,A)}, \quad \sigma_3(A, E, \delta) = \tfrac{2\pi}{T_2(E,A)}. \tag{4.11.57}$$

This follows after some contemplation of Eqs. (4.11.42) and (4.11.51).

Again we do not enter into the analysis of the regularity and invertibility of the integration map $I(\dot{\overline{\theta}}, \dot{\overline{\varphi}} \dot{\overline{\psi}}, \overline{\theta}, \overline{\varphi}, \overline{\pi}) \longleftrightarrow (\delta, E, A, \gamma, \widetilde{\varphi}, \widetilde{\psi})$.

Note that the coordinates chosen in the general case do not reduce to those of the symmetric case $(I_1 = I_2)$ when $I_2 \to I_1$.

However, there is great arbitrariness in defining the prime integrals because any function of δ, E, A is still a prime integral, and it is possible to find two other prime integrals Φ, Ψ becoming $\dot{\varphi}$ and $\dot{\psi}$ in the $I_1 = I_2$ case. In fact, let

$$\Phi = \frac{1}{T_1(E, A)} \int_0^{T_1(E,A)} A \frac{I_1 \Omega_1(t)^2 + I_2 \Omega(t)^2}{I_1^2 \Omega_1(t)^2 + I_2^2 \Omega(t)^2} \, dt, \tag{4.11.58}$$

where $\Omega(t)$ is defined in Eq. (4.11.30) and $\Omega_1(t)$ is related to Ω by Eq. (4.11.29) with $\Omega(t)$ replacing $\Omega_3(t)$ [and, likewise, we could define $\Omega_3(t)$ by Eq. (4.11.30)].

Note that Φ is the average value of $\dot\varphi$ along a period [since $\dot\varphi$ is periodic with period $T_1(E,A)$, see Eq. (4.11.47)].

Analogously, from Eq. (4.11.12), (4.11.13) written without overbars, one can find an expression of $\dot\psi$ in terms of $\omega_1,\omega_2,\omega_3$:

$$\dot\psi = \frac{(A^2-2EI_3)\omega_3}{I_1^2\omega_1^2+I_2\omega_2^2} \tag{4.11.59}$$

So $\dot\psi$ is a periodic function with period $T_1(E,A)$ and we can define the prime integral

$$\Psi = \frac{1}{T_1(E,A)}\int_0^{T_1(E,A)} \frac{(A^2-2EI_3)\Omega_3(t)}{I_1^2\Omega_1(t)^2+I_2\Omega(t)^2} \tag{4.11.60}$$

where the ambiguity of the sign in the definition of Ω_3 has, now, to be resolved by remarking that ω_3 from Eq. (4.11.30) never vanishes if $A\neq 0$ and, therefore, it has a constant sign which we attribute also to Ω_3.

It could also be possible to change $\widetilde\psi$ to a variable reducing to ψ, when $I_2\to I_1$. However, we shall not do this.

It remains to check Eq. (4.11.54); $T_2=\frac{2\pi}{\chi_0}$:

$$\chi_0(E,A) = \frac{2}{T_1(E,A)}\int_0^{\frac12 T_1(E,A)} \Phi(t,E,A)dt \tag{4.11.61}$$

and changing variable $t\to\Omega(t,E,A)$, one has [see Eq. (4.11.43)]

$$dt = \frac{d\Omega}{\sqrt{-V_{E,A}(\Omega)}} \tag{4.11.62}$$

Hence, recalling that Ω_1, can be expressed in terms of Ω, we can express the integral on the right-hand side of Eq. (4.11.61) as an integral over the variable Ω, via Eqs. (4.11.48) and (4.11.29) and, after some algebra, Eq. (4.11.54) follows. mbe

4.11.1 Problems and Complements

1. Let $\widetilde{\mathcal{L}}$ be the Lagrangian function describing the motion of a rigid body in a fixed frame $(O;\bar{\mathbf{i}},\bar{\mathbf{j}},\bar{\mathbf{k}})$ in Euler angle coordinates

$$\widetilde{\mathcal{L}} = \frac12 I_1(\dot{\overline\theta}\cos\overline\psi+\dot{\overline\varphi}\sin\overline\theta\sin\overline\psi)^2+\frac12 I_2(-\dot{\overline\theta}\sin\overline\psi+\dot{\overline\varphi}\sin\overline\theta\cos\overline\psi)^2+\frac12 I_3(\dot{\overline\varphi}\cos\overline\theta+\dot{\overline\psi})^2$$

Compute the canonical variables $p_{\overline\theta},p_{\overline\varphi},p_{\overline\psi}$ associated with $\overline\theta,\overline\varphi,\overline\psi$ via $\widetilde{\mathcal{L}}$. Show that if $\mathbf{K}_O$ is the angular momentum of the solid with respect to the fixed point O, and if $\overline{\mathbf{n}}$ is the node line unit vector, then

$$p_{\overline\theta}=\mathbf{K}_O\cdot\overline{\mathbf{n}},\quad p_{\overline\varphi}=\mathbf{K}_O\cdot\overline{\mathbf{k}},\quad p_{\overline\psi}=\mathbf{K}_O\cdot\mathbf{i}_3.$$

(*Hint:* Just apply the definition of p [Eq. (3.11.1)] and use Eqs. (4.11.11) and (3.9.3).)

2. Call $A = |\mathbf{K}_O|$, $K_z = p_{\overline{\varphi}}$, $L = \mathbf{K}_O \cdot \mathbf{i}_3 = p_{\overline{\psi}}$ and let (γ, φ, π) be the angles considered on p.318 (see Fig. 4.10). $(K_z, A, L, \gamma, \varphi, \psi)$ are the "Deprit variables" ([13]).

$$\overline{\mathbf{n}} \stackrel{def}{=} (\bar{\mathbf{i}}, \bar{\mathbf{j}}) \cap (\mathbf{i}_1, \mathbf{i}_2), \qquad \mathbf{n} \stackrel{def}{=} (\mathbf{i}_1, \mathbf{i}_2) \cap (\mathbf{i}, \mathbf{j}), \qquad \mathbf{m} \stackrel{def}{=} (\bar{\mathbf{i}}, \bar{\mathbf{j}}) \cap (\mathbf{i}, \mathbf{j}) \equiv \mathbf{i}$$

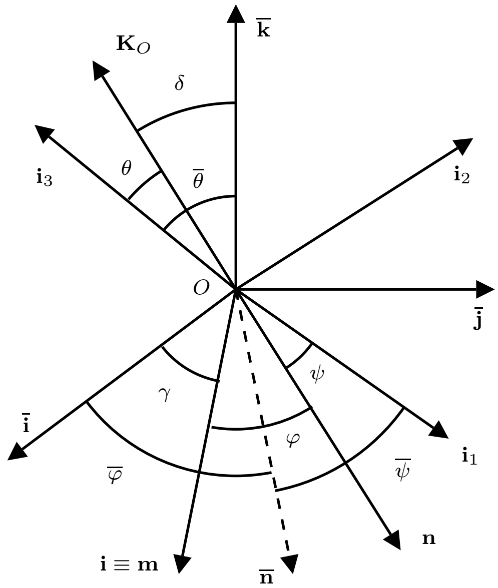


Figure 4.10: The Deprit angles. Here $\overline{\mathbf{n}}$ is the node line $(\bar{\mathbf{i}}, \bar{\mathbf{j}}) \cap (\mathbf{i}_1, \mathbf{i}_2)$, $\mathbf{n}$ is the node line $(\mathbf{i}_1, \mathbf{i}_2) \cap (\mathbf{i}, \mathbf{j})$ and $\mathbf{m} \equiv \mathbf{i}$ is the node $(\mathbf{i}, \mathbf{j}) \cap (\bar{\mathbf{i}}, \bar{\mathbf{j}})$. The $\mathbf{j}$ axis is not drawn.

Show that given $(p_{\overline{\theta}}, p_{\overline{\varphi}}, p_{\overline{\psi}}, \overline{\theta}, \overline{\varphi}, \overline{\psi})$, the Deprit variables are determined and vice versa. (*Hint:* Note that $p_{\overline{\varphi}} = A \cos\delta = K_z$, $p_{\overline{\psi}} = A \cos\theta = L$, $p_{\overline{\theta}} = -A \sin\theta \, \sin(\psi - \overline{\psi})$, and note that the angles $\varphi, \theta, \overline{\psi} - \psi, \overline{\varphi} - \gamma, \delta, \overline{\theta}$ can be arranged in a spherical triangle (Fig. 4.11).

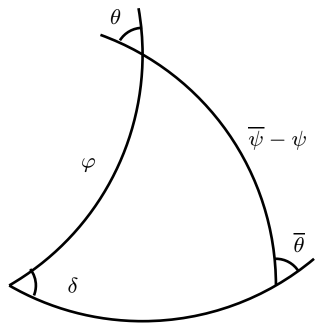


Figure 4.11.The spherical triangle associated with the Deprit's angles.

Therefore, given the Deprit variables, one computes $p_{\overline{\varphi}} = K_z$, then $\cos\delta = \frac{K_z}{A}$, then $p_{\overline{\psi}} = L$, then $\cos\theta = \frac{L}{A}$. Hence, at this point, one knows the elements φ, θ, δ of the spherical

triangle in Fig. 4.11 and by solving it one computes, by spherical trigonometry, the three other elements, i.e., $\overline{\psi}-\psi, \overline{\varphi}-\gamma, \overline{\theta}$, and since γ, ψ are known, one gets $\overline{\psi}$. Consider the spherical triangle of Fig. 4.12.

3. Consider the spherical triangle of Fig. 4.12. Check the basic spherical trigonometry relations:

$$\begin{aligned}
&(1) \qquad \cos C = \cos A \cos B + \sin A \sin B \cos\gamma \\
&(2) \qquad \cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos C \\
&(3) \qquad \frac{\sin\alpha}{\sin A} = \frac{\sin\beta}{\sin B} = \frac{\sin\gamma}{\sin C} \\
&(4) \qquad \sin C\cos\beta = \cos B\sin A - \sin B\cos A\cos\gamma \\
&(5) \qquad \cos A\cos\gamma = \sin A\cot B - \sin\gamma\cot\beta \\
&(6) \qquad dA = \cos\beta dC + \cos\gamma dB + \sin B\sin\gamma d\alpha
\end{aligned}$$

(*Hint:* Draw the spherical triangle in Fig. 4-12 by locating the vertex 2 with the angle γ on the z axis, the vertex 1 with the β angle on the xz plane: so that the three vertices are expressed in Cartesian coordinates as $\mathbf{r}_1 = (\sin A, 0, \cos A)$, $\mathbf{r}_2 = (0,0,1)$ and $\mathbf{r}_3 = (\sin B\cos\gamma, \sin B\sin\gamma, \cos B)$. Then
to check (1) note that $\mathbf{r}_1\cdot\mathbf{r}_3 = \cos C$;
to check (2) apply (1) to the spherical triangle formed on the sphere by the perpendicular to the planes containing the arcs A, B, C;
to check (3) note that $\mathbf{r}_1\cdot\mathbf{r}_2\wedge\mathbf{r}_3 = \sin A\sin B\sin\gamma$ has to be symmetric in the interchange of the role of $(A,\alpha),(B,\beta),(C,\gamma)$;
to check (4) remark that $\mathbf{r}_1\wedge\mathbf{r}_3\cdot\mathbf{j} = -\sin C\cos\beta$;
the identity (5) is a consequence of (1) and (4);)

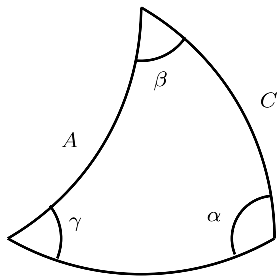


Figure 4-12: Spherical triangle with the sides formed by the arcs A, B, C opposite to the angles α, β, γ.

4. Show that the map $(p_{\overline{\theta}}, p_{\overline{\varphi}}, p_{\overline{\psi}}, \overline{\theta}, \overline{\varphi}, \overline{\psi}) \longleftrightarrow (K_z, A, L, \gamma, \varphi, \psi)$ has the property ("Deprit's theorem"):

$$K_z d\gamma + A d\varphi + L d\psi = p_{\overline{\theta}} d\overline{\theta} + p_{\overline{\varphi}} d\overline{\varphi} + p_{\overline{\pi}} d\overline{\psi}.$$

(*Hint:* Using Problems 2 and 3, show that $d\varphi = \cos\theta d(\overline{\psi}-\psi) + \cos\delta d(\overline{\varphi}-\gamma) - \sin\theta\sin(\overline{\psi}-\psi)d\overline{\theta}$; then substitute into the left-hand side using $K_z = p_{\overline{\varphi}}$, $L = p_{\overline{\psi}}$ and $-A\sin\theta\sin(\overline{\psi}-\psi) = p_{\overline{\theta}}$.)

5. The map $(p_{\overline{\theta}}, p_{\overline{\varphi}}, p_{\overline{\psi}}, \overline{\theta}, \overline{\varphi}, \overline{\psi}) \longleftrightarrow (K_z, A, L, \gamma, \varphi, \psi)$, defined in Problems 2 and 4 maps six variables into six others without any reference to a rigid body. Interpret Problem 4 as saying

that this map is a completely canonical map homogeneous in the variables in the sense of Proposition 22, §3.11, p.224. (*Hint:* Apply Proposition 22, §3.11.)

6. Compute the rigid body's Hamiltonian $\widetilde{H}$ in Deprit variables, remarking that, by Problem 5, it must simply be the kinetic energy expressed in these variables (see the general properties of the completely canonical transformations, §3.12), and show that

$$\widetilde{H}(K_z, A, L, \gamma, \varphi, \psi) = \frac{1}{2}\frac{L^2}{I_3} + \frac{1}{2}\Big(\frac{\sin^2\psi}{I_1} + \frac{\cos^2\psi}{I_2}\Big)(A^2 - L^2).$$

Deduce the Hamilton equations of the motion and check that they are identical to Eqs. (4.11.47) and (4.11.59). Use this Hamiltonian formulation to rederive directly the integrability of the motions of a solid with a fixed point. (*Hint:* Note that the kinetic energy can be derived from $\mathbf{K}_O = (\sqrt{A^2 - L^2}\sin\psi, \sqrt{A^2 - L^2}\cos\psi, L)$. Write the equations of motion and integrate by quadratures.)

7. Using the Hamiltonian in Problem 6, show that the solid with a fixed point gives rise to canonically integrable motions (see Definition 11, §4.8, p.289). (*Hint:* Since the map $(p_{\overline{\theta}}, p_{\overline{\varphi}}, p_{\overline{\psi}}, \overline{\theta}, \overline{\varphi}, \overline{\psi}) \longleftrightarrow (K_z, A, L, \gamma, \varphi, \psi)$ is completely canonical, it is enough to show that the Hamiltonian motions generated by the Hamiltonian in Problem 6 are canonically integrable. The $\widetilde{H}$ has a ψ dependence and, at the same time, it also involves A: but one just finds the canonical transformation $(L, \psi) \longleftrightarrow (M, \mu)$ that integrates the 1-degree of freedom system in which A' is considered a parameter and keeps track of the obvious implications on the other variables. The procedure is standard and it is discussed as an example. Define the canonical transformation with generating function

$$\Phi(K'_z, A', M, \gamma, \varphi, \psi) = K'_z\gamma + A'\varphi + S(A', M, \psi)$$

with S chosen so that Φ solves the Hamilton-Jacobi equation for $\widetilde{H}$:

$$\frac{1}{2}\Big(\frac{1}{I_3} - \frac{\cos^2\psi}{I_2}\frac{\sin^2\psi}{I_1}\Big)\Big(\frac{\partial S}{\partial\psi}\Big)^2 + \frac{A'^2}{2}\Big(\frac{\sin^2\psi}{I_1} + \frac{\cos^2\psi}{I_2}\Big) = e(A', M)$$

where the function $e(A', M)$ is naturally chosen so that the function S does generate a canonical transformation on the Hamiltonian $\widetilde{H}$, regarded as a function of L, ψ only (parameterized by $A' \equiv A$), bringing it to action angle variables (M, μ). By Problem 5, §3.11, this means that the function $e(A', M)$ has to be chosen so that

$$\frac{\partial e(A', M)}{\partial M} = \omega(A', E)$$

where $\omega(A', E)$ is the pulsation of the motion (of this one-dimensional system parameterized by A') with energy $E = e(A', M)$. Since the equation of motion for ψ in this auxiliary one-dimensional system is

$$\dot{\psi} = -\frac{\partial\widetilde{H}}{\partial L}(A', L, \psi),$$

the pulsation will be such that

$$\frac{2\pi}{\omega(A', E)} = \int_0^{2\pi}\frac{d\psi}{\dot{\psi}(t)} = \int_0^{2\pi}\frac{d\psi}{-\frac{\partial\widetilde{H}}{\partial L}(A', L)},$$

where L has to be fixed so that $\widetilde{H}(A', L, \psi) = E$; i.e., L has to be taken as a function $L(E, A', \psi)$:

$$L = L(E, A', \psi) = \sqrt{2\frac{E - A'^2\left[\frac{\cos^2\psi}{I_2} + \frac{\sin^2\psi}{I_3}\right]}{\frac{1}{I_3} - \frac{\cos^2\psi}{I_2} - \frac{\sin^2\psi}{I_3}}}$$

which permits us to compute $\omega(A', E)$.

The function $e(A', M)$ can be computed in terms of its inverse $m(A', E)$ (such that $m(A', e(A', M)) \equiv M$), since $\frac{\partial m}{\partial E})(A', E)$ must be

$$\Big(\frac{\partial e}{\partial M}(A', M)\Big)^{-1} = \frac{1}{\omega(A', E)}.$$

So, for instance, e can be defined by inverting the relation:

$$m(A', E) = \int_{E_0(A')}^{E} \frac{dE'}{\omega(A', E)},$$

where $E_0(A') = \min_{L,\psi} \widetilde{H}(A', L, \psi)$.

Coming back to S, we see that

$$S(A', M, \psi) = \int_0^{\psi} L(e(A', M), A', \psi')\, d\psi'$$

is the explicit solution of the Hamilton-Jacobi equation (recall the expression of L).

The above Φ-generated canonical transformation leaves A, K_z, γ unchanged and changes L to M, φ to some new φ' and ψ to some new μ with

$$\varphi' = \varphi + \frac{\partial S}{\partial A}(A, M, \psi), \qquad \mu = \frac{\partial S}{\partial M}(A, M, \psi)$$

and transforms the $\widetilde{H}$ into $e(A, M)$.

The above transformation is "globally" defined because one can show that

$$S(A', M, 2\pi) \equiv 2M.$$

In fact, $\frac{\partial S}{\partial M}(A, M, 2\pi) \equiv 2\pi$ (since this can be checked directly by differentiating the integral giving S and by comparing it to the integral for computing $\omega(A, E)$ explicitly and then using $\frac{\partial e}{\partial M} = \frac{1}{\omega(A,E)}$; ; so $S(A', M, 2\pi) = 2\pi M + g(A')$ for some function $g(A')$. But $M = 0$ means $E = E_0(A')$; hence, $L = 0$; hence, $S \equiv 0$; hence, $g(A') \equiv 0$. This means that when (th, φ) vary on $\mathcal{T}^2$, $\varphi', \mu)$ also vary on $\mathcal{T}^2$.)

The following problems provide a simple example of how to use the canonical formalism for a concrete application. A more complete treatment of the problem will be presented in Ch. 5, as an application of perturbation theory.

8. (*Solar precession Hamiltonian*) Imagine that the Earth $\mathcal{E}$ is an ideally rigid homogeneous solid of rotation with equatorial radius R. Assume that the center T revolves on a purely Keplerian orbit $t \to \mathbf{r}_T(t)$ and, see Fig.4.10, fix the frame $\overline{\mathbf{i}}, \overline{\mathbf{j}}, \overline{\mathbf{k}}$ to be with center T and with $\overline{\mathbf{k}}$ axis orthogonal to the plane of the Earth orbit, while the $\overline{\mathbf{i}}$ axis is at the equinox line at a prefixed time (*epoch*). Show that the motion of the Earth is described in the coordinates $(\overline{\theta}, \overline{\varphi}, \overline{\psi})$ of problem (1) above, by the Lagrangian:

$$\mathcal{L} = \frac{1}{2}J(\dot{\overline{\varphi}}\cos\overline{\theta} + \dot{\overline{\psi}})^2 + \frac{1}{2}I(\dot{\overline{\theta}}^2 + \dot{\overline{\varphi}}^2 \sin^2\overline{\theta}) + \int_{\mathcal{E}} \frac{kM_S}{|\mathbf{r}_T + \mathbf{x}|}\frac{d\mathbf{x}}{|\mathcal{E}|}$$

with $J = I_3$, $I = I_1 = I_2$ being the Earth inertia moments, M_T, M_S being the masses of the Earth and of the Sun, k being the gravitational constant and $|\mathcal{E}|$ being the Earth volume: in

the case of an ellipsoid with polar radius $(1-\eta)R$ it is $J = (2/5)R^2 M_T$, $I = J(1-\eta+\eta^2/2)$. (Hint: show that I, I_3 are the appropriate inertia moments and remark that in the given geocentric frame of reference the axes have a fixed orientation; hence the inertia forces (constant per unit mass and due only to the drag, as the Coriolis force vanishes) have vanishing moment with respect to T, by symmetry. Hence in the chosen comoving frame we have an ideal solid body subject to the gravitational attraction whose potential, in a configuration respecting the constraint of rigidity, is precisely the above integral).

9. Show that the integral in the Lagrangian of problem (8) can be written:

$$-V = C_1(t) + \frac{3kM_S M_T}{2} \int_{\mathcal{E}} \frac{(\mathbf{r}_T \cdot \mathbf{x})^2}{|\mathbf{r}_T|^2} \frac{1}{|\mathbf{r}_T|^3} \frac{d\mathbf{x}}{|\mathcal{E}|} + O((\frac{R}{|\mathbf{r}_T|})^4)$$

where $C_1(t)$ is a suitable function of t. (*Hint:* By Taylor expansion: $|\mathbf{r}_T + \mathbf{x}|^{-1} = |\mathbf{r}_T|^{-1} - |\mathbf{r}_T|^{-2}(\mathbf{r}_T \cdot \mathbf{x})/|\mathbf{r}_T| + (3(\mathbf{r}_T \cdot \mathbf{x})^2/|\mathbf{r}_T|^2 - \mathbf{x}^2)/2|\mathbf{r}_T|^3 + O((R/|\mathbf{r}_T)^3)$; developing to fourth order one sees that the third order also vanishes.)

10. Show that V in problem (9) can be written:

$$-V = C_2(t) + \frac{3kM_S}{2|\mathbf{r}_T|^3} \frac{I-J}{J} J \cos^2\alpha = C_2(t) - \frac{3}{2} \frac{kM_S}{|\mathbf{r}_T|^3} \eta_1 J \cos^2\alpha$$

where $C_2(t)$ is a suitable function, α_S= angle between the symmetry axis $\bar{\mathbf{i}}_3$ and the vector $\mathbf{r}_T$, and $\eta_1 \equiv (J-I)/I$. (*Hint:* just compute explicitly the integrals over $\mathbf{x}$ in problem (9).)

11. Let $\bar{\mathbf{i}}$ be as in proble (8). Then the angle $\overline{\varphi}$ is called the *precession* angle since the equinox fixing epoch. If the Earth longitude (*i.e.* the angle between the position $\mathbf{r}_T$ and $\bar{\mathbf{i}}$) is λ_T, then the *apparent longitude* is $\lambda_T - \overline{\varphi}$. Show that:

$$\cos\alpha = -\sin\overline{\theta}\sin(\lambda_T - \overline{\varphi})$$

(Hint: write $\bar{\mathbf{i}}_3 = (\sin\overline{\theta}\sin\overline{\varphi}, -\sin\overline{\theta}\cos\overline{\varphi}, \cos\overline{\theta})$ and $\mathbf{r}_T/|\mathbf{r}_T| = (\cos\lambda_T, \sin\lambda_T, 0)$, in the geocentric frame, and compute the scalar product). (*Hint:* Compute explicitly the integrals over $\mathbf{x}$ in (9).)

12. Using fig.4.11, 4.12, 4.10 and the trigonometric relations for general spherical triangles in problem (3), plus the second of the following other identities of spherical trigonometry:

$$\sin C\cos\beta = \cos B\sin A - \sin B\cos A\cos\gamma$$
$$\cos A\cos\gamma = \sin A\cot B - \sin\gamma\cot\beta$$

show that the inversion in (2) can be actually performed via the relations:

$$\cos\delta = \frac{K_z}{A}, \qquad \cos\theta = \frac{L}{A}$$
$$\cot(\overline{\varphi} - \gamma) = (\cos\varphi\cos\delta + \sin\delta\cot\theta)/\sin\varphi$$
$$\cot(\overline{\psi} - \psi) = (\cos\varphi\cos\theta + \sin\theta\cot\delta)/\sin\varphi$$
$$\sin\overline{\theta} = \sin\theta\frac{\sin\varphi}{\sin(\overline{\varphi} - \gamma)}$$

(Remark: to check the two spherical identities in problem (3) and the two above simply draw the spherical triangle putting the vertex 2 with the angle γ on the z axis, the vertex 1 with the β angle on the xz plane so that the three vertices are expressed in cartesian coordinates as $\mathbf{r}_1 = (\cos A, 0, \sin A)$, $\mathbf{r}_2 = (0,0,1)$ and $\mathbf{r}_3 = (\sin B\cos\gamma, \sin B\sin\gamma, \cos B)$.

Then observe that $\mathbf{r}_1\cdot\mathbf{r}_3=\cos C$, that $\mathbf{r}_1\cdot\mathbf{r}_2\wedge\mathbf{r}_3$ has to be symmetric in the interchange of the role of $(A,\alpha),(B,\beta),(C,\gamma)$ and that $\mathbf{r}_1\wedge\mathbf{r}_3\cdot\mathbf{j}=-\sin C\cos\beta$; the three latter relations, after computing the left hand sides in cartesian coordinates for $\mathbf{r}_i$ yield, respectively, the two identities in problem (3) and the first of the above two; the last identity is a consequence of the first and third.)

13. Using the coordinates $(K_z,A,L,\gamma,\varphi,\psi)$ show that the Hamiltonian describing the above system for the theory of the solar precession is:

$$\frac{1}{2}(\frac{L^2}{J}+\frac{A^2-L^2}{I})+\frac{3kM_S}{2|\mathbf{r}_T|^3}\eta_1 J\cos^2\alpha_S+O((\frac{R}{|\mathbf{r}_T|})^4)$$

and, setting $\eta_1\equiv(J-I)/J$, $\eta_2=(J-I)/I$ and neglecting $O((R/|\mathbf{r}_T|)^4)$, it becomes:

$$H_p=\frac{1}{2}\frac{A^2}{J}+\eta_2\Big(\frac{A^2-L^2}{J}\Big)+\eta_1\frac{3kM_S}{2|\mathbf{r}_T|^3}J\cos^2\alpha_S$$

so that in the case of an ellipsoidal Earth: $\eta_1=\eta-\eta^2/2$, $\eta_2=\eta+\eta^2/2+O(\eta^4)$. (*Hint:* By problems (10),(11) the term V added to the Lagrangian depends only on the coordinates $(\overline{\theta},\overline{\varphi},\overline{\psi})$: hence the conjugate momenta $p_{\overline{\theta}},p_{\overline{\varphi}},p_{\overline{\psi}}$ are given by the same expression as when $V=0$, see problem (1). Therefore in the $(p_{\overline{\theta}},p_{\overline{\varphi}},p_{\overline{\psi}},\overline{\theta},\overline{\varphi},\overline{\psi})$ variables the hamiltonian is simply the same hamiltonian with $V=0$ plus V expressed in terms of the new variables and of time. Finally the map $(p_{\overline{\theta}},p_{\overline{\varphi}},p_{\overline{\psi}},\overline{\theta},\overline{\varphi},\overline{\psi})\to(K_z,A,L,\gamma,\varphi,\psi)$ is a completely canonical time independent map; hence the hamiltonian in the last variables is just the old one evaluated in the new variables.)

14. Using proble (12) show that the Hamiltonian H_p can be written:

$$\begin{aligned}H_p=&\frac{1}{2}\frac{A^2}{J}+\eta_2\frac{A^2-L^2}{2J}+\eta_1\Big(\frac{3kM_S}{2|\mathbf{r}_T|^3}J\cdot\\&\cdot[\sin(\lambda_T-\gamma)\Big(\cos\varphi\sin\theta\cos\delta+\sin\delta\cos\theta\Big)-\cos(\lambda_T-\gamma)\sin\theta\sin\varphi]^2\Big)=\\=&\frac{1}{2}\frac{A^2}{J}+\eta_2\frac{A^2-L^2}{2J}+\eta_1\Big(\frac{3kM_S}{2|\mathbf{r}_T|^3}J\cdot\\&\cdot[\sin(\lambda_T-\gamma)\Big(\frac{K_z}{A}(1-\frac{L^2}{A^2})^{1/2}\cos\varphi+\frac{L}{A}(1-\frac{K_z^2}{A^2})^{1/2}\Big)\\&-(1-\frac{L^2}{A^2})^{1/2}\cos(\lambda_T-\gamma)\sin\varphi]^2\Big)\end{aligned}$$

15. Suppose that the excentricity of the Earth orbit is neglected (*i.e.* that the orbit of the Earth is taken circular with radius a equal to the major semiaxis of the ellipse), show that the average over the angles φ,γ and over time t of H_p is:

$$\overline{H}_p=\frac{A^2}{2J}+\eta_2\frac{A^2-L^2}{2J}+\eta_1\Big(\frac{3kM_S}{2a^3}J[\frac{K_z^2}{A^2}(1-\frac{L^2}{A^2})\frac{1}{4}+\frac{1}{2}(1-\frac{K_z^2}{A^2})\frac{L^2}{A^2}+\frac{1}{4}(1-\frac{L^2}{A^2})]\Big)$$

if a is the major semiaxis of the Earth orbit, with an error of order $O(\eta e)$. The latter hamiltonian describes the motion ove time scales large compared to those if the slowest period in the non averaged Hamiltonian.

16. Suppose that $A=L$ (*i.e.* neglect the non alignment between the Earth axis and the angular momentum), so that $\gamma=\overline{\varphi}$. And, furthermore, assume that the hamiltonian H_p can be replaced by $\overline{H}_p$ for the purpose of evaluating the average motion over many periods of

revolutions (see Ch. 5, §10/12 for a more rigorous treatment). Then show that the precession angular velocity would be:

$$\dot{\gamma} = \lambda_p^S = -\eta_1 \frac{3kM_S}{2a^3} \frac{JK_z}{2A^2} + O(\eta e^2)$$

In this approximation the angles δ, θ are constant and, having negleted θ, δ has the interpretation of inclination angle i_0 of the Earth axis. Using the Kepler's law: $kM_S/a^2 = \omega_T a$, if $\omega_T = 2\pi/T$ is the angular velocity of the mean anomaly and T is the period of the Earth revolutionand if ω_D is the angular velocity of the daily rotation, show that the solar precession rate is:

$$\lambda_p^S = -\frac{3}{2}\eta_1 \omega_T^2 \frac{JK_z}{A^2} = -\frac{3}{2}\eta_1 \frac{\omega_T}{\omega_D}\omega_T \cos i_0$$

the fact that is negative is often referred as a *retrograde* precession. Show also that the period of precession, is $T_p^S = -2\pi/\lambda_p^S = T(\omega_D T \cos i_0)/3\pi\eta$, or since $T = 1.$ year and $\eta = 0.00335281$, $T_p^S \sim 7.94\,10^4$years. (*Hint:* $A = I_3\omega_D$, in the suggested approximation, and $K_z = A\cos i_0$. Use then the connection (4.10.18), *i.e.* the third Kepler's law, between the Earth axis a, the period T and the gravitational constant kM_S: the relation $A = I_3/\omega_D$ is correct only to a first approximation, evaluate the exact value, still neglection the θ, and check that this is really neglegible. Also the relation between the period and the gravitational constant is correct if we neglect the ratio of the masses M_T/M_S: check that if we do not want to neglect it the correction would be an extra factor $(1 + M_T/M_S)$).

17. A rough analysis of the lunar precession can be made assuming that the Moon is on the ecliptic and that its orbit is circular. Show that the solar precession analysis can the be applied to the Moon influence and that the lunar precession would be, if M_L, a_L denote respectively the Moon mass and the radius of its orbit:

$$\dot{\gamma} = \lambda_p^L = -\eta_1 \frac{3kM_T}{2a_L^3} \frac{JK_z}{A^2} + O(\eta_1 e^2) = \lambda_p^S (\frac{a}{a_L})^3 \frac{M_L}{M_S}$$

so that the total luni-solar precession would be:

$$\lambda_p = \lambda_p^S + \lambda_p^L = \lambda_p^S \Big(1 + (\frac{a}{a_L})^3 \frac{M_L}{M_S}\Big) \sim 3\lambda_p^S$$

Evaluate the total rate of lunisolar precession in the above approximation and show that it gives $T_p \sim 2.51\,10^4$ years (get the data from appendix P), or a yearly precession of the equinoxes of $\sim 51.6''$ per year. So that only 1/3 of the luni-solar precession is due to the Sun. Show that even assuming that Jupiter gravitated around the Earth on a circular orbit its contribution to the precession would be much smaller (Hint: with obvious notations it would be a fraction of the order of $(a/a_J)^3 M_J/M_S$, *i.e.* $O(10^{-5})$ of the solar precession).

The observed value of the lunisolar precession is however $50.38''$ per year: the discrepancy is due to the crudeness of the approximations in the model. A more accurate calculation (Laplace) leads to a formula which was in fact *used* to determine η from the known precession rate, in terms of the masses of the Sun and of the Moon. Check that corrections come from several sources:
(1) the eccentricity of the Moon orbit, the inclination of the Moon orbit and the eccentricity of the Earth orbit have been neglected.
(2) the center of mass of the Earth-Moon rather than that of the Earth revolve about the Sun on a keplerian orbit.

18. Correct the above theory for the eccentricity of the Earth. This means that, looking for the motion on scales of time large with respect to T, *i.e.* the period of revolution, we

do not write $(a/|\mathbf{r}_T|)^3 \cos^2 \alpha_S$ using the approximations $\lambda_T = \omega_T t + const$ and $a/|\mathbf{r}_T| = 1$, but we use the Kepler laws (see problem (13), p.304), *i.e.*:

$$\begin{aligned} |\mathbf{r}_T| =& a(1 + e\cos\xi) \\ \xi =& \lambda - e\sin\lambda + (e^2/2)\sin 2\lambda \\ \lambda_T =& \lambda - 2e\sin\lambda + (5/4)\sin 2\lambda \\ \lambda =& \omega_T t + const \end{aligned}$$

and then evaluate the average of $(a/|\mathbf{r}_T|)^3 \cos^2 \alpha_S$ over φ and λ, still neglecting θ, *i.e.* $(1 - L^2/A^2)$. Show that the result is:

$$\begin{aligned} &\big\langle (1 + e\cos\xi)^{-3} \sin^2(\lambda_T - \gamma)\big\rangle = (1 - 4e^2) \\ &H = \frac{A^2}{2I_3} + \eta_2 \frac{A^2 - L^2}{2I_3} + \eta_1 \frac{3}{2} I_3 \omega_T^2 \Big(1 - \frac{K^2}{A^2}\Big)(1 - \frac{3}{2}e^2) \end{aligned}$$

19. Correct the Moon contribution to take into account the eccentritcity e_L and the inclination i_L of the Moon orbit. Calling $\theta_L, \varphi_L, \psi_L$ he Euler angles of the frame with $\mathbf{j}_L$ axis orthogonal to the Moon orbit, $\mathbf{n}_L$ the node of the Moon orbit with the ecliptic, x axis, $\mathbf{i}math_L$ pointing to the actual position of the Moon, and if α_L is the angle between the Moon Earth axis $\mathbf{r}_L$ and the Earth axis, it is:

$$-\cos\alpha_L = \cos\psi_L \sin\overline{\theta}\sin(\varphi_L - \overline{\varphi}) + \cos\psi_L \sin\overline{\theta}\cos(\varphi_L - \overline{\varphi}) + \cos\psi_L - \sin\theta_L \cos\overline{\theta}\sin\psi_L$$

and then chech that the average $\big\langle (a_L/|\mathbf{r}_L|)^3 \cos^2\alpha_L \big\rangle$ is, still neglecting θ is:

$$(1 + \frac{3}{2}e_L^2)(1 - \frac{3}{2}\sin^2\theta_L)$$

(*Hint:* Write the coordinates of $\mathbf{r}_L$ and $\imath_3$ in the fixed frame as:

$$\begin{aligned} \imath_3 =& (-\sin\overline{\theta}\cos\overline{\varphi}, -\sin\overline{\theta}\cos\overline{\varphi}, \cos\overline{\theta}) \\ \mathbf{r}_L/|\mathbf{r}_L| =& (-\sin\psi_L \cos\theta_L \sin\varphi_L + \cos\psi_L \cos\varphi_L, \\ & \sin\psi_L \cos\theta_L \cos\varphi_L + \cos\psi_L \sin\varphi_L,\ \sin\theta_L \sin\psi) \end{aligned}$$

and then use that the motion of the Moon is Keplerian for ψ_L and a uniform precession for φ_L).

20. Put togheter the last problems to check that the lunisolar precession is given by:

$$-\omega_p = \lambda^s \Big((1 + \frac{3}{2}e_T^2) + \frac{M_S}{M_L}\Big(\frac{a}{a_L}\Big)^3 (1 + \frac{3}{2}e_L^2)(1 - \frac{3}{2}\sin^2 i_L) \Big)$$

and check that the data in appendix Q give $-\omega_p = 51.51$” per year. Further corrections can be found by avoiding to take averages over time scales of the order of the year (theory of *nutation*).

4.12 Integrable Systems. Geodesic Motion on the Surface of an Ellipsoid and Other Systems

In general, given a closed regular surface $\Sigma \subset \mathcal{R}^d$, the “geodesic motions” are the motions which a unit mass point can undergo on Σ, when it is ideally bound to Σ and subject to no other active forces.

The Lagrangian of such motions is

$$\mathcal{L} = \frac{1}{2}\dot{\mathbf{x}}^2 \tag{4.12.1}$$

and energy conservation, §3.5, implies that

$$T = \frac{1}{2}\dot{\mathbf{x}}^2 = \text{constant} \tag{4.12.2}$$

on the considered motions. For instance, the set of the motions with initial speed of modulus 1 consists entirely of motions in which the speed has modulus 1 at all times.

The "geodesic flow" on Σ is the flow on $\mathcal{S} = \{$data space for the motions on $\Sigma\} = \{$set of pairs $(\boldsymbol{\eta}, \mathbf{x})$ with $\mathbf{x} \in \Sigma$ and $\boldsymbol{\eta}$ compatible with the constraint, i.e., tangent to $\Sigma\}$[11] which to every point $(\boldsymbol{\eta}, \mathbf{x}) \in \mathcal{S}$ associates $S_t(\boldsymbol{\eta}, \mathbf{x}) \in \mathcal{S}$, the configuration into which the datum $(\boldsymbol{\eta}, \mathbf{x})$ evolves in time t under the only influence of an ideal constraint to Σ.

Since $|\dot{\mathbf{x}}|$ = constant, there is an a priori bound on the distance that a point can travel in a given time and, therefore, the geodesic flow is well defined, $\forall\, t \in \mathcal{R}$.

Speed conservation has an interesting consequence: the action of a geodesic motion $t \to \mathbf{x}(t)$ computed between t_1 and t_2 can be expressed in terms of the curvilinear abscissas on the trajectory on which $\mathbf{x}$ moves. If $V = |\dot{\mathbf{x}}|$,

$$A_{t_1,t_2} = \{\text{action of } \mathbf{x} \text{ between } t_1 \text{ and } t_2\} = \frac{V^2\,(t_2 - t_1)}{2} = \frac{V}{2}(s_2 - s_1) \tag{4.12.3}$$

By the least-action principle, Proposition 8, §3.5, p.163, we know that the motion $t \to \mathbf{x}(t)$ makes the action locally minimal in sufficiently small time intervals.

From this, it follows that the trajectory $\mathcal{I}$, as a curve in $\mathcal{R}^d$, makes the distance between $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$ measured along Σ locally minimal if t_2 is close enough to t_1 ("Maupertuis' principle", see problems to §3.11). In fact, given $\mathbf{x}_1 = \mathbf{x}(t_1) \in \mathcal{I}$, suppose that for $|t_2 - t_1| < \varepsilon$, the action $A_{t_1 t_2}(\mathbf{y})$ is minimal on Σ as $\mathbf{y}$ varies in $\mathcal{M}_{t_1,t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); \Sigma) = \{$motions on Σ defined for $t \in [t_1, t_2]$ and leading from $\mathbf{x}(t_1)$ to $\mathbf{x}(t_2)\}$. If there existed a curve $\mathcal{C}_{1,2}$ connecting $\mathbf{x}(t_1)$ with $\mathbf{x}(t_2)$, lying on Σ and shorter than $(s_2 - s_1) = \{$length of the part of $\mathcal{I}$ between $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)\}$, then one could run it with uniform speed starting from $\mathbf{x}(t_1)$ at time t_1 so as to reach $\mathbf{x}(t_2)$ at time t_2.

Such a motion $\mathbf{x}_{\mathcal{C}_{1,2}} \in \mathcal{M}_{t_1,t_2}(\mathbf{x}(t_1), \mathbf{x}(t_2); \Sigma)$ would have an action

$$\begin{aligned} A_{t_1,t_2}(\mathbf{x}_{\mathcal{C}_{1,2}}) =& \frac{1}{2}\Big(\frac{|\mathcal{C}_{1,2}|}{t_2 - t_1}\Big)^2 (t_2 - t_1) = \frac{1}{2}\frac{|\mathcal{C}_{1,2}|^2}{t_2 - t_1} \\ <& \frac{1}{2}(s_2 - s_1)^2 t_2 - t_1 = \frac{1}{2}V^2(t_2 - t_1) = \frac{V}{2}(s_2 - s_1) \end{aligned} \tag{4.12.4}$$

[11] In fancy language, call this the "tangent fiber bundle" to Σ.

as $|\mathcal{C}_{1,2}| < s_2 - s_1 = \{\text{length of } \mathcal{I}\}$. This contradicts the minimality of A_{t_1,t_2} on $\mathbf{x}$.

The curves on a surface Σ which make minimal the distance between the points that they connect provided such points are close enough, are called "geodesics" on Σ, and this explains the name given to the motions with Lagrangian (4.12.1) on Σ.

The simplest nontrivial example of a geodesic motion is the motion on the surface of the sphere in $\mathcal{R}^3$. The possible trajectories of this motion are great circles. It is possible to interpret this statement in terms of the integrability of the geodesic motion on the surface of the sphere in the sense of Definition 10, §4.8, p.287. In this case, the motions are all periodic (see Observation (5), p.288).

A less simple example is the motion on the surface of the ellipsoid. We shall only treat the case of the ellipsoid of revolution. However, the motion on an arbitrary ellipsoid is also integrable (see problems at the end of this section for a glimpse of the theory).

In the case of an ellipsoid of revolution, we choose as z-axis the symmetry axis of the ellipsoid $\mathcal{E}$ and determine the position on $\mathcal{E}$ of a point through the two coordinates (θ, φ) as

$$x = a\sin\theta\,\cos\varphi, \quad y = a\sin\theta\,\sin\varphi, \quad z = b\cos\theta, \tag{4.12.5}$$

where a and b are the principal semi-axes of the ellipsoid. The Lagrangian (4.12.1) of the geodesic motion on $\mathcal{E}$ can be written by Eq. (4.1.5) as

$$\mathcal{L}(\dot\theta, \dot\varphi, \theta, \varphi) = \frac{1}{2}[(b^2\sin^2\theta + a^2\cos^2\theta)\,\dot\theta^2 + a^2\dot\varphi^2\sin^2\theta]. \tag{4.12.6}$$

So that the equations of motion are

$$\frac{d}{dt}a^2\,\dot\varphi\,\sin^2\theta = 0, \tag{4.12.7}$$

$$\frac{d}{dt}(b^2\sin^2\theta + a^2\cos^2\theta)\dot\theta = \frac{\partial\mathcal{L}}{\partial\theta}(\dot\theta, \dot\varphi, \theta, \varphi) \tag{4.12.8}$$

However, it is convenient to discuss only Eq. (4.12.7), combining it with the energy conservation principle:

$$\frac{1}{2}[(b^2\sin^2\theta + a^2\cos^2\theta)\,\dot\theta^2 + a^2\dot\varphi^2\sin^2\theta] = E \tag{4.12.9}$$

Equations (4.12.9) and (4.12.7), which we use to define the prime integral $A = \dot\varphi\,\sin^2\theta$, yield

$$\dot\theta = \pm\sqrt{\frac{2E\sin^2\theta - a^2A^2}{\sin^2\theta\,(b^2\sin^2\theta + a^2\cos^2\theta)}} \stackrel{def}{=} \pm\sqrt{-V_{E,A}(\theta)} \tag{4.12.10}$$

which, by the usual argument, implies that $t \to \theta(t)$ is periodic with period:

$$T_1(E,A) = 2\int_{\theta_-(E,A)}^{\theta_+(E,A)} \frac{d\theta}{\sqrt{-V_{E,A}(\theta)}}, \tag{4.12.11}$$

where $\theta_-(E,A)$ and $\theta_+(E,A)$ are the two solutions of $V_{E,A} = 0$ of the form $\theta_\pm(E,A) = \frac{\pi}{2} \pm \theta_0(E,A)$ or $\theta_(E,A) = -\frac{\pi}{2} \pm \theta_0(E,A)$ and $\theta_0 = \arcsin(\frac{aA}{2E})$. Furthermore, θ verifies the equation

$$\ddot{\theta} = -\frac{\partial V_{E,A}}{\partial \theta}(\theta) \tag{4.12.12}$$

and, therefore, it is a C^∞ function of t (see §2.7) and can be expressed in terms of the solution $t \to R(t,E,A)$ of Eq. (4.12.12) with initial data $R(0,E,A) = \theta_-(E,A), \dot{R}(0,E,A) = 0$. Such a function is defined, recalling §2.7, by

$$t = \int_{\theta_-(E,A)}^{R(t,E,A)} \frac{d\theta}{\sqrt{-V_{E,A}(\theta)}}, \tag{4.12.13}$$

for $0 \le t < \frac{T_1(E,A)}{2}$; it is continued naturally for $\frac{T_1(E,A)}{2} < t < T_1(E,A)$. Furthermore, $\theta(t)$ is given by

$$\theta(t) = R(t + t_0(\theta_0, \dot{\theta}_0), E, A), \tag{4.12.14}$$

where $t_0(\theta_0, \dot{\theta}_0) =$ {first time when the motion $t \to R(t,E,A)$ reaches $(\dot{\theta}_0, \theta_0)$, with $\theta_0 = \theta(0)$, $\dot{\theta}(0)$.

As one sees, the analysis of this problem by "quadratures" is entirely analogous to the ones seen in §4.9-§4.11. As on those occasions, the motion $t \to \varphi(t)$ can be deduced from Eq. (4.12.7) by a quadrature,

$$\varphi(t) = \varphi_0 + \int_0^t \frac{A}{\sin^2\theta(t)} dt \tag{4.12.15}$$

It can be treated, as already seen in §4.9-4.11, by noting that

$$\frac{A}{\sin^2 R(t,E,A)} = \sum_{k=-\infty}^{\infty} \chi_k(A,E) e^{\frac{2\pi\, i\, k}{T_1(E,A)} t} \tag{4.12.16}$$

by the periodicity of R and by Fourier's theorem, as $t \to \frac{A}{\sin^2 R(t,E,A)}$ is a $T_1(E,A)$-periodic C^∞-function with Fourier coefficients $(\chi_k(A,E))_{k\in\mathcal{Z}}$. Setting

$$S(t,E,A) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{+\infty} \chi_k(A,E) \frac{e^{\frac{2\pi\, i\, k}{T_1(E,A)} t}}{\frac{2\pi i k}{T_1(E,A)}} \tag{4.12.17}$$

the quadrature (4.12.15) yields

$$\varphi(t) = \varphi_0 + \chi_0(E,A)t + S(t + t_0(\theta_0, \dot\theta_0, E, A) - S(t_0(\theta_0, \dot\theta_0), E, A), \quad (4.12.18)$$

Hence, from Eqs. (4.12.14) and (4.12.18), we can conclude that all motions are quasi-periodic with periods $T_1(E,A)$ given by Eq. (4.12.11) and

$$T_2(E,A) = \frac{2\pi}{\chi_0(A,E)} = \frac{T_1(E,A)}{\int_{\theta_-(E,A)}^{\theta_+(E,A)} \frac{A}{\sin^2\theta} \frac{d\theta}{\sqrt{-V_{E,A}(\theta)}}} \qquad (4.12.19)$$

after changing variables $\theta = R(t,E,A)$ along the lines already seen in Proposition 24, §4.11, p.314, and Proposition 22, §4.9, p.293.

It could be checked that as E, A vary the two periods $T_1(E,A)$, $T_2(E,A)$ will generally have an irrational ratio.

The above analysis basically achieves the proof of the following proposition, (if one disregards the checks of regularity and invertibility in suitably large regions W' in the data space of the map $I(\dot\theta(0), \dot\varphi(0), \theta(0), \varphi(0)) = (E, A, \alpha, \beta)$ with $\alpha = \frac{2\pi}{T_1(E,A)} t_0(\theta(0), \dot\theta(0))$, $\beta = \varphi(0) - S(t_0(\theta(0), \dot\theta(0)), E, A)$:

25 Proposition. *The set W of the data for the geodesic motions on an ellipsoid of revolution $\mathcal{E}$ and such that $E \neq 0, A \neq 0$ can be covered by sets $W' \subset W$ on which the motions are are integrable in the sense of Definition 10, §4.8, p.287. Such motions are quasi-periodic with periods $T_1(E,A), T_2(E,A)$, given by Eqs. (4.12.11) and (4.12.19). If the ellipsoid semi-axes are different the motion is generally quasi periodic and non periodic.*

Observations.
(1) The discussion preceding Proposition 25 is very general and could be repeated with essentially no change to cover very general classes of surfaces of revolution like those parametrically described by equations like

$$z = f(\theta), \qquad x = g(\theta)\cos\varphi, \qquad y = g(\theta)\sin\varphi \qquad (4.12.20)$$

for $(\theta, \varphi) \in [0, 2\pi] \times [0, 2\pi]$, with $f, g, \in C^\infty(\mathcal{T}^1)$ such that the curve in $\mathcal{R}^2$ with parametric equations $\xi = g(\theta)$, $\eta = f(\theta)$, $\theta \in [0, 2\pi]$ is a simple closed curve *symmetric* under reflection around the η axis.

Other surfaces covered by the above method are those with parametric equations

$$z = a(\varphi), \qquad x = b(\varphi)\cos\psi, \qquad y = b(\varphi)\sin\psi \qquad (4.12.21)$$

for $(\varphi, \psi) \in [0, 2\pi] \times [0, 2\pi]$, with $a, b \in C^\infty(\mathcal{T}^1)$ such that the curve in $\mathcal{R}^2$ with parametric equations $\eta = a(\varphi), \xi = b(\varphi)$, $\varphi \in [0, 2\pi]$ is a simple closed curve contained in the half-plane $\xi > 0$. The reader can check the above statements, as an exercise on the quadrature method.
(2) Surfaces like Eq. (4.12.20) generalize the ellipsoid of revolution while those like Eq. (4.12.21) generalize the "torus of revolution": given $a, b > 0, a > b$,

$$x = (a + b\,\cos\varphi)\cos\psi, \qquad y = (a + b\,\cos\varphi)\sin\psi, \qquad z = b\,\sin\varphi. \quad (4.12.22)$$

We conclude this list of remarkable integrable systems by citing a few other systems integrable on suitable regions W.
(1) A point mass on an ellipsoid of revolution, with symmetry axis along the major axis of the revolving ellipse, subject to a force with potential energy

$$V(x) = \frac{g}{|\mathbf{x} - \mathbf{f}_1| \cdot |\mathbf{x} - \mathbf{f}_2|}, \qquad (4.12.23)$$

where $\mathbf{f}_1, \mathbf{f}_2$ are the foci of the ellipse generating the ellipsoid.

This system can be integrated by the quadrature method of §4.9-§4.11, and one obtains similar results in elliptic coordinates defined in terms of Cartesian coordinates (x, y, z) of $\mathbf{x}$ as

$$\sqrt{x^2 + y^2} = \sigma\,\sqrt{(\xi^2 - 1)(1 - \eta^2)}, \quad z = \sigma\,\xi\,\eta, \quad \text{azimuth of} = \varphi \quad (4.12.24)$$

where $\xi \in [1, +\infty], \eta \in [-1, 1]$, $\varphi \in [0, 2\pi]$ and the parameter σ has to be chosen so that the considered ellipsoid is a $\xi =$ *constant* surface. Such surfaces are the ellipsoids

$$\frac{z^2}{\sigma^2\xi^2} + \frac{x^2 + y^2}{\sigma^2(\xi^2 - 1)} = 1. \qquad (4.12.25)$$

(2) A unit mass on a sphere with potential energy in polar coordinates:

$$U(\theta, \varphi) = b(\theta) + \frac{c(\varphi)}{\sin^2\theta} \qquad (4.12.26)$$

with b, c C^∞-periodic functions with period 2π. This system is integrated by quadratures by writing its Lagrangian function in polar coordinates and discussing the Lagrange equations.
(3) A solid body with a symmetry axis fixed at a point 0 of this axis, which we call $\mathbf{i}_3$, different from the center of mass G and subject to ideal constraints plus the weight, i.e., a force $m_i\mathbf{g}$ on the i-th point (equivalent by Observations (5) in §3.2, p.148, to a force Mg, $M = \sum_i m_i$ applied to G as far as the force momentum calculation is concerned).
This system ("heavy gyroscope") is also integrable by quadratures: proceeding as in §3.11, choose the fixed reference frame $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ with $\mathbf{k}$ axis anti parallel to $\mathbf{g}$, and write the Lagrangian function in terms of the Euler angles. The Lagrange equations can then be combined with the conservation laws, for energy and for the $\mathbf{k}$ component of the angular momentum, to reduce the problem to that of the analysis of one-dimensional systems, i.e., to the quadratures. See, also, problems at the end of this section and problems to §3.5 in [28]).
(4) Two more difficult classical integrable systems are the geodesic motions on the surface of a non symmetric ellipsoid (see problems at the end of this

section for an introduction to this theory) and the motions of a heavy rigid body with a fixed point O, with the baricenter in the $\mathbf{i}_1\mathbf{i}_2$ plane, say on the $\mathbf{i}_1$ axis at distance a from O, and with inertia moments $I_1 = I_2 = 2I_3 \stackrel{def}{=} 2I$. Such systems can be shown to be integrable by quadratures (as discovered by Jacobi and Kovalevskaya, respectively, see problems).
(5) Other systems are N point masses on the line $\mathcal{R}$ with Lagrangian functions

$$\mathcal{L} = \frac{m}{2}\sum_{i=1}^{N}\dot{x}_i^2 - g\sum_{i=1}^{N-1} e^{\alpha(x_i - x_{i+1})} \tag{4.12.27}$$

called the "Toda lattice", or

$$\mathcal{L} = \frac{m}{2}\sum_{i=1}^{N}\dot{x}_i^2 - g^2\sum_{i<j}\frac{1}{(x_i-x_j)^2} - \frac{\omega^2}{2}\sum_{i=1}^{N}x_i^2, \tag{4.12.28}$$

called "Calogero lattice", respectively. These were discovered very recently and are also integrable. Some variants of such systems with the same properties are also known.
(6) Obviously, there are other integrable systems: it suffices to perform an arbitrary change of coordinates in the Lagrangian functions which we have just examined to obtain Lagrangian functions of integrable.
However, only very "few" other systems are known that have the integrability property and that are "interesting", i.e., not obtained by trivial changes of coordinates from those so far listed. Some can be found among the problems for §4.12.
Finally, we remark that all the integrable systems of §4.9-§4.12 could be shown to be not only integrable in the sense of Definition 10, §4.8, p.287, but also analytically and canonically integrable in the sense of Definition 11, §4.8, p.289, in large regions of the phase space. In the problems of §4.10-4.12, the main steps towards such a proof are given.

4.12.1 Exercises and Problems

1. Integrate explicitly by quadratures the systems mentioned in the points (1), (2), and (3) of the list of integrable systems in §4.12. By the Hamilton-Jacobi method, show their canonical integrability.

2. Integrate the heavy gyroscope system (3) p.331, by using the Deprit variables (see problems to §4.11). First show that the Hamiltonian (i.e., the energy) can be written in the Deprit variables as $H(K_z, A, L, \gamma, \varphi, \psi)$ given by

$$\frac{A^2-L^2}{2I} + \frac{L^2}{2I_3} + \mu\Big[\frac{K_z L}{A^2} - \sqrt{1-\frac{K_z^2}{A^2}}\sqrt{1-\frac{L^2}{A^2}}\cos\psi\Big],$$

where $\mu = Mgd$, $M =$ total mass, $g =$ gravity constant, $I \equiv I_1 = I_2$ and I_3 are the moments of inertia. Show also the canonical integrability of this system.

3. Consider the "Kovalevskaya gyroscope", see p.331, and show that its Lagrangian is

$$\mathcal{L} = I\dot{\theta}^2 + I\sin^2\theta\,\dot{\varphi}^2 + \frac{I}{2}(\dot{\psi} + \dot{\varphi}\cos\theta)^2 + Mga\sin\theta\cos\psi$$

and explicitly write the Lagrange equations relative to the θ, φ, ψ variables.

4. In the context of Exercise 3, eliminate the $\dot{\psi}$ between the Lagrange equations relative to φ and ψ and add the resulting equation to the equation relative to the θ variable multiplied by $+i$ or $-i$, $(i = \sqrt{-1})$, successively. Show that the two resulting equations imply $\frac{1}{U}\frac{dU}{dt} = -\frac{1}{V}\frac{dV}{dt} =''$ same function$''$, where:

$$U = (\dot{\varphi}\,\sin\theta + i\,\dot{\theta})^2 + Mgae^{-i\psi}\sin\theta = \overline{V},$$

so $UV \equiv |U|^2 =$ constant and UV is a prime integral.[12]

5. In the context of Problems 3 and 4, show that the Kovalevskaya gyroscope is integrable by quadratures on vast regions of phase space.

6. Consider the geodesic motion on the surface of the ellipsoid $\mathcal{E}$: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, $a < b < c$. Introduce the local coordinate system ("Jacobi's system") described by

$$x = \sqrt{a\tfrac{(u-a)(v-a)}{(b-a)(c-a)}}, \quad y = \sqrt{b\tfrac{(u-b)(v-b)}{(c-b)(a-b)}}, \quad z = \sqrt{c\tfrac{(u-c)(v-c)}{(a-c)(b-c)}},$$

for $(u, v) \in [b, c] \times [a, b]$ or $(u, v) \in [a, b] \times [b, c]$. Defining for $\lambda \in \mathcal{R}$

$$A(\lambda) = \frac{1}{4}\frac{\lambda}{(a-\lambda)(b-\lambda)(c-\lambda)},$$

show that the kinetic energy is given by

$$T(\dot{u}, \dot{v}, u, v) = \frac{1}{2}(u - v)\big(A(u)\dot{u}^2 - A(v)\dot{v}^2\big).$$

Applying the Hamilton-Jacobi method to the Lagrangian system with Lagrangian $\mathcal{L} = T(\dot{u}, \dot{v}, u, v)$, show that the geodesic motion on the ellipsoid admits a second prime integral:

$$M(\dot{u}, \dot{v}, u, v) = (u - v)(vA(u)\dot{u}^2 - uA(v)\dot{v}^2).$$

(*Hint:* Write the Hamilton-Jacobi equation in (u, v) variables after finding the Hamiltonian function in (u, v) and in their canonically conjugate momenta p_u, p_v:

$$\frac{\partial f}{\partial t} + \frac{1}{2}\beta\Big[\big(\frac{\partial f}{\partial u}\big)^2\frac{1}{H(u,v)} + \big(\frac{\partial f}{\partial v}\big)^2\frac{1}{G(u,v)}\Big] = 0,$$

where $H(u, v) = (u - v)A(u)$, $G(u, v) = (v - u)A(v)$, and look for solutions of the form

$$f(u, v, t) = -\frac{E}{2}t + \psi(u, v), \qquad \psi(u, v) = \alpha(u) + \beta(v)$$

The equation becomes

$$A(v)\big(\frac{\partial \psi}{\partial u}\big)^2 - A(u)\big(\frac{\partial \psi}{\partial v}\big)^2 = (u - v)\,A(u)\,A(v)\,E,$$

admitting a family of solutions parameterized by E and a new arbitrary parameter a:

---

[12] See [49], Chap. VI.

$$\psi(u,v|a,E) = \int_{u_0}^{u} \sqrt{EA(u')(u'+a)}du' + \int_{v_0}^{v} \sqrt{EA(v')(v'+a)}dv'.$$

Now, applying the canonical transformation G generated by $-\frac{E}{2}t + \psi(u,v,a,E)$, deduce that the trajectories of the motion (geodesics on $\mathcal{E}$) are given by the equation $\frac{d\psi}{da} = c = cnstant$, i.e., $F_a(u) + F_a(v) = 2c$ if $F_a(u)$ is a primitive function to $\sqrt{A(u)(u+a)}$. This also implies that a is a prime integral. Writing the canonical transformation G, it is possible to express a in terms of u, v, p_u, p_v or $u, v, \dot{u}, \dot{v}$. The computation gives $a = -M(\dot{u}, \dot{v}, u, v)/T(\dot{u}, \dot{v}, u, v) = -M/E$; so M is a prime integral. ([10]).)

7. Consider the system ("atom in electric field")

$$H(\mathbf{p}, \mathbf{q}) = \frac{\mathbf{p}^2}{2m} - \frac{g}{|\mathbf{q}|} + Fx$$

$\mathbf{p} = (p_x, p_y, p_z)$, $\mathbf{q} = (x, y, z)$ and study it in "squared parabolic coordinates"

$$x = \frac{1}{2}(u^2 - v^2), \quad y = uv\cos\varphi, z = uv\sin\varphi$$

and show by the method of problem 6 (i.e., by the Hamilton-Jacobi method) that this system has three prime integrals and that it can be integrated by quadratures (from [46]).

8. Consider the Hamiltonian ("ionized hydrogen molecule")

$$H(\mathbf{p}, \mathbf{q}) = \frac{\mathbf{p}^2}{2m} - \frac{g}{|\mathbf{q} - \mathbf{f}_1|} - \frac{g}{|\mathbf{q} - \mathbf{f}_2|}$$

with $\mathbf{p} = (p_x, p_y, p_z) \in \mathcal{R}^3$, $\mathbf{q} = (x, y, z) \in \mathcal{R}^3$, $\mathbf{f}_1, \mathbf{f}_2 \in \mathcal{R}^3$ and study it in elliptical coordinates (see Eq. (4.12.24) and (4.12.25)) and show by the methods of problems 6 and 7 that it has three prime integrals and that it can be integrated by quadratures. Find canonical action-angle variables (from [46]).

4.13 Some Integrability Criteria. Introduction: Geometric Considerations and Preliminary Definitions

Considering the "rarity" of the mechanical systems known as integrable one wonders whether it is possible to easily recognize, a priori, the non integrability of a mechanical system.

For instance, the integrability on a region W of the data space $\mathcal{S}$ implies the existence of ℓ-"independent" prime integrals. Therefore, a way of showing non integrability might be that of showing the nonexistence of as many prime integrals as the number of degrees of freedom.

In any concrete case, however, it is very difficult to decide whether or not a system possesses prime integrals (other than the total energy and its functions). Poincaré's proof of non integrability, in a sense stricter than the above, of the motion of three heavenly bodies is based on showing the nonexistence

of enough prime integrals (also defined in a more stringent way). It is still a famous proof (see [38], vol. 1, ch. VI).

Hence, it is useful to try to identify other special properties of the integrable systems to use them as necessary conditions for integrability or to formulate sufficient non integrability conditions.

In the following sections we go through an analysis that will allow us to classify the motions of the integrable systems as “simple and ordered” motions and those of the non integrable ones as “complex and disordered”.

Coming back into the frame of mind of §2.21 and §4.8, the notions of observable, average values, etc. introduced there have a natural extension to the systems with several degrees of freedom.

We consider an ℓ-degrees-of-freedom system described by a Lagrangian function

$$\mathcal{L}(\dot{\mathbf{x}}, \mathbf{x}) \ + \ \textit{ideal constraints} \tag{4.13.1}$$

regular in the sense of §3.11, Definition 14, and generating Hamiltonian equations admitting global solutions, in the future and in the past for all the constraint-compatible initial data.

As usual, we denote $\mathcal{S}$ the data space for the system of Eq. (4.13.1). By Proposition 18, p.285, it is a regular surface in $\mathcal{R}^d \times \mathcal{R}^d$ where d is the dimension of the unconstrained system, usually $d = 3N$, $N = \{$number of points in the system$\}$, see Definition 9, §4.8, p.287.

12 Definition. *The elements of $C^\infty(\mathcal{S})$[13] will be called the “observables” of the mechanical system of Eq. (4.13.1).*
Given an increasing sequence $\mathbf{t} = t_0, t_1, \ldots$ such that $t_i \xrightarrow[i\to+\infty]{} +\infty$ and given $f \in C^\infty(\mathcal{S})$, we shall call the “$\mathbf{t}$-history of f” on the motion of $(\dot{\mathbf{x}}, \mathbf{x}) \in \mathcal{S}$ the sequence

$$(f(S_{t_i}(\dot{\mathbf{x}}, \mathbf{x})))_{i=0}^{\infty} \tag{4.13.2}$$

It is the sequence of the results of the successive observations of the values of f on the motion starting at $(\dot{\mathbf{x}}, \mathbf{x})$ at times $t_0, t_1 \ldots$. We shorten the notation by simply referring to the “$(f, \mathbf{t})$-history of $(\dot{\mathbf{x}}, \mathbf{x})$”.

If $f \in C^\infty(\mathcal{S})$ and $\mathbf{t}$ is a sequence like

$$t_i = i\, t_1, \qquad i = 0, 1, \ldots; t_1 > 0 \tag{4.13.3}$$

and if $(\dot{\mathbf{x}}, \mathbf{x}) \in W \subset \mathcal{S}$, where W is a region on which the mechanical system of Eq. (4.13.1) is integrable, then the $(f, \mathbf{t})$-history of $(\dot{\mathbf{x}}, \mathbf{x})$ is far from being an “arbitrary” sequence of numbers. Proposition 6,§4.2, p.251, allows us to state, for instance, the following obvious reformulation of its contents.

26 Proposition. *If in $W \subset \mathcal{S}$ the system of Eq. (4.13.1) is integrable and $f \in C^\infty(\mathcal{S})$ is an observable and if $\mathbf{t}$ is as in Eq. (4.13.3), the $(f, \mathbf{t})$ histories*

[13] See Observation (2) to Definition 7, p.285.

of the points $(\dot{\mathbf{x}}, \mathbf{x}) \in W$ *have a well-defined average value, i.e., the following limit exists:*

$$\overline{f}(\dot{\mathbf{x}}, \mathbf{x}) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} f(S_{jt_1}(\dot{\mathbf{x}}, \mathbf{x})). \tag{4.13.4}$$

Furthermore, if $(\mathbf{A}, \boldsymbol{\varphi}) = I(\dot{\mathbf{x}}, \mathbf{x})$ *is the integrating transformation mapping* W *onto* $V \times \mathcal{T}^\ell$*, see Definition 10, §4.8, p.287, and if the* $(\ell + 1)$ *numbers* $(\omega_1(\mathbf{A}), \ldots, \omega_\ell(\mathbf{A}), \sigma)$*, with* $\sigma = \frac{2\pi}{t_1}$ *are rationally independent then*

$$\overline{f}(\dot{\mathbf{x}}, \mathbf{x}) = \frac{1}{(2\pi)^\ell} \int_{\mathcal{T}^\ell} F_f(\mathbf{A}, \boldsymbol{\varphi}')\, d\boldsymbol{\varphi}', \tag{4.13.5}$$

having set

$$F_f(\mathbf{A}, \boldsymbol{\varphi}) = f(I^{-1}(\mathbf{A}, \boldsymbol{\varphi})). \tag{4.13.6}$$

Observations.
(1) F_f is the observable f in the new $(\mathbf{A}, \boldsymbol{\varphi})$ coordinates.
(2) Hence, the non integrability of the system of Eq. (4.13.1) in W can be proved by "just" exhibiting a single point $(\dot{\mathbf{x}}, \mathbf{x}) \in W$ and a single observable f whose $(f, \mathbf{t})$ history on $(\dot{\mathbf{x}}, \mathbf{x})$ does not have a well-defined average value.
(3) However, this criterion is very difficult to apply in practice: the $(f, \mathbf{t})$ histories are very hard to analyze in concrete interesting cases and "usually" they admit an average value even in non integrable systems.

The following proposition provides a more geometric integrability criterion different in spirit from the one above.

27 Proposition. *If in* $W \subset \mathcal{S}$ *the system of Eq. (4.13.1) is integrable, the closure of every trajectory of points* $(\dot{\mathbf{x}}, \mathbf{x}) \in W$ *is a set* $\overline{\mathcal{I}}$ *which can be mapped continuously and in a one-to-one way onto a torus* $\mathcal{T}^s$ *with* $1 \le s \le \ell$*, if* ℓ *is the number of degrees of freedom of the system.*

Observations.
(1) Proposition 27 is also essentially a way of rephrasing some properties of the integrability of Definition 10, §4.8, p.287. In fact the motions of an integrable system take place on invariant tori of dimension ℓ run quasi-periodically. If $(\omega_1, \ldots, \omega_\ell)$ are the pulsations of a given motion and are rationally independent then the trajectory fills densely a set homeomorphic to $\mathcal{T}^\ell$, [14] see Proposition 4, p.250, §4.2. In general, if s is the number of elements of a maximal subset of $(\omega_1, \ldots, \omega_\ell)$ consisting of rationally independent numbers, then $\overline{\mathcal{I}}$ will be homeomorphic to $\mathcal{T}^s$. The proof of this fact is left to the reader and is essentially described in the hints to the problems for §4.14.
(2) So to prove non integrability in W, it suffices to find "just" one $(\dot{\mathbf{x}}, \mathbf{x}) \in W$

[14] i.e., a set which is a one-to-one bicontinuous image of $\mathcal{T}^\ell$.

whose trajectory has a closure which is not homeomorphic to a smooth surface or, more particularly, to an s-dimensional torus, $s \le \ell$.
(3) The geometric structure of a trajectory of a point mass bound to a surface Σ can be found via Maupertuis' principle since the trajectory of an energy-E motion is a geodesic for the metric $dh = \sqrt{2(E - V(\boldsymbol{\xi}))}ds$ on Σ, where ds is the line element on Σ and V is the potential energy of the active forces.
In geometry, some criteria for the existence of dense geodesics on a bounded surface with some metric are known (e.g., if the curvature of the metric is everywhere negative, there are dense geodesics). So with the help of Proposition 27 and of the Maupertuis' principle, some examples of non integrable systems can be easily built.

To obtain deeper insight into integrable systems, it is convenient to restrict attention to the "analytically integrable" systems.

They are connected with some interesting geometrical notions which we have to illustrate before continuing the analysis.

To help the reader avoid getting lost in the labyrinth of the geometric concepts that follow, it is better to state our aim at the beginning. Basically we wish to define sets $G \subset \mathcal{R}^d \times \mathcal{T}^d$ with "piecewise analytic boundary" (see the following definition of analyticity). Such sets have the remarkable property that not only are they measurable in the Riemann sense, but also that their intersection with planar surfaces are measurable with respect to the Riemann measure on the surface. This is a property which might not hold for sets with C^∞ boundary (see Problems).

We shall need, in an essential way, the above simple property and its invariance with respect to some changes of coordinates. There are several ways of constructing families of sets and classes of coordinate changes with this property. However, none of them seems describable in few words, although this fact might seem surprising. It will be an amusing puzzle for the reader to try to find (possibly giving up analyticity) some alternative definitions which would allow us to retain the substance of §4.14 and §4.15.

13 Definition. *If $\Omega \subset \mathcal{R}^d$ is open and $f \in C^\infty(\Omega)$, then f is "analytic" on Ω if $\forall \boldsymbol{\xi}_0 \in \Omega$, $\exists \varepsilon(\boldsymbol{\xi}_0) > 0$ such that f can be developed $\forall |\boldsymbol{\xi} - \boldsymbol{\xi}_0| < \varepsilon(\boldsymbol{\xi}_0)$, as*

$$f(\boldsymbol{\xi}) = \sum_{k_1,\ldots,k_d}^{0,\infty} \frac{\partial^{k_1+\ldots+k_d} f}{\partial \xi_1^{k_1} \ldots \partial \xi_d^{k_d}}(\boldsymbol{\xi}_0) \prod_{=1}^{d} \frac{(\xi_j - (\boldsymbol{\xi}_0)_j)^{k_j}}{k_j!} \qquad \text{and} \tag{4.13.7}$$

$$\sum_{k_1,\ldots,k_d}^{0,\infty} \Big| \frac{\partial^{k_1+\ldots+k_d} f}{\partial \xi_1^{k_1} \ldots \partial \xi_d^{k_d}}(\boldsymbol{\xi}_0) \Big| \frac{\varepsilon(\boldsymbol{\xi}_0)^{k_1+\ldots+k_d}}{k_1! \ldots k_d!} < +\infty. \tag{4.13.8}$$

If $\mathbf{k} = (k_1, \ldots, k_d) \in \mathcal{Z}^d$, define $\mathbf{k}! = \prod_{j=1}^d k_j!$, $(\boldsymbol{\xi} - \boldsymbol{\xi}_0)^{\mathbf{k}} = \prod_{j=1}^d (\xi_j - (\boldsymbol{\xi}_0)_j)^{k_j}$ and $\partial^{\mathbf{k}} f(\boldsymbol{\xi})$ for $\frac{\partial^{k_1+\ldots+k_d} f}{\partial \xi_1^{k_1} \ldots \partial \xi_d^{k_d}}(\boldsymbol{\xi})$, so that Eq. (4.13.7) will be rewritten as

$$f(\boldsymbol{\xi}) = \sum_{\mathbf{k}\in \mathcal{Z}_+^d} \frac{\partial^{\mathbf{k}} f(\boldsymbol{\xi}_0)}{\mathbf{k}!}(\boldsymbol{\xi}-\boldsymbol{\xi}_0)^{\mathbf{k}}, \tag{4.13.9}$$

An $\mathcal{R}^d$-valued function on Ω is called analytic if its components are analytic.

In the following, we will also need the obvious extension of the notion of analytic function f to the case when Ω is an open subset in $\mathcal{R}^d \times \mathcal{T}^{d'}$ and its values are in $\mathcal{R}^\ell \times \mathcal{T}^{\ell'}$, $d+d' > 0$, $\ell+\ell' > 0$, $d, d', \ell, \ell' > 0$.

First remark that a real function on $\Omega \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ can be "canonically extended" to a function f on a set $\widetilde{\Omega} \subset \mathcal{R}^d \times \mathcal{R}^{d'}$ by setting

$$\begin{aligned} &\widetilde{\Omega} = \{\text{pairs } (\boldsymbol{\xi},\boldsymbol{\eta}) \in \mathcal{R}^d \times \mathcal{R}^{d'} \text{ with } (\boldsymbol{\xi},\boldsymbol{\eta} \bmod 2\pi) \equiv (\boldsymbol{\xi},\boldsymbol{\varphi}) \in \Omega\} \\ &\widetilde{f}(\boldsymbol{\xi},\boldsymbol{\eta}) \stackrel{def}{=} f(\boldsymbol{\xi},\boldsymbol{\eta} \bmod 2\pi) \equiv f(\boldsymbol{\xi},\boldsymbol{\varphi}). \end{aligned} \tag{4.13.10}$$

With the above convention (4.13.10) and with the above restrictions on d, d', ℓ, ℓ', we state the following definition (for some examples see Exercises).

14 Definition. *A function on $\Omega \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ taking values in $\mathcal{R}^\ell$ and associating to $\boldsymbol{\xi} \in \Omega$ the value $(\mathbf{x},\boldsymbol{\varphi})$ will be called "analytic" on Ω if, $\forall\, \boldsymbol{\xi}_0 \in \Omega$, there is a function F, "representative of f", defined in the vicinity of $\boldsymbol{\xi}_0$, taking values in $\mathcal{R}^\ell \times \mathcal{R}^{\ell'}$ and analytic, such that*

$$\text{if } F(\boldsymbol{\xi}) = (\mathbf{x},\boldsymbol{\eta}) \text{ then } f(\boldsymbol{\xi}) = (\mathbf{x},\boldsymbol{\varphi}) \text{ with } \boldsymbol{\varphi} = \boldsymbol{\eta} \bmod 2\pi \tag{4.13.11}$$

for all $\boldsymbol{\xi}$ near $\boldsymbol{\xi}_0$.
A function f on an open set $\Omega \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ taking values in $\mathcal{R}^\ell \times \mathcal{T}^{\ell'}$ will be called analytic on Ω if its canonical extension $\widetilde{f}$ to $\widetilde{\Omega}$ is analytic.
The derivatives of f will obviously be defined as the "derivatives of the canonical extension of a representative" and they will be denoted by the usual symbols.

Observation. If some of the integers ℓ, ℓ', d, d' vanish, we interpret $\mathcal{R}^\ell \times \mathcal{T}^{\ell'}$ or $\mathcal{R}^d \times \mathcal{T}^{s'}$ in the obvious way: $\mathcal{R}^0 \times \mathcal{T}^p \equiv \mathcal{T}^p$, $\mathcal{R}^p \times \mathcal{T}^0 \equiv \mathcal{R}^p$, $\forall\, p > 0$.

Together with the notion of analytic function, we need the notion of analytic coordinates.

15 Definition. *Let $U \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ be open and let $\boldsymbol{\Xi}$ be an $\mathcal{R}^d \times \mathcal{T}^{d'}$-valued analytic function defined on an open set $\Omega \subset \mathcal{R}^d \times \mathcal{R}^{d'}$ such that:*
(i) $\boldsymbol{\Xi}$ is invertible as a map between U and Ω;
(ii) the Jacobian determinant of $\boldsymbol{\Xi}$ never vanishes on Ω ("$\boldsymbol{\Xi}$ is nonsingular");[15]
(iii) $\boldsymbol{\Xi}$ and $\boldsymbol{\Xi}^{-1}$ are analytic in Ω and U, respectively. Then we say that $(U,\boldsymbol{\Xi})$ is an analytic system of local coordinates on U.

[15] Naturally, the Jacobian determinant of $\boldsymbol{\Xi}$ in $\boldsymbol{\xi}_0$ is the Jacobian determinant of a representative of $\boldsymbol{\Xi}$ near $\boldsymbol{\xi}_0$ (see Definition 14, above).

If $U \subset \mathcal{R}^d \times \mathcal{T}^{d'}$, $V \subset \mathcal{R}^{\overline{d}} \times \mathcal{T}^{\overline{d}'}$ are open sets and $d + d' = \overline{d} + \overline{d}'$, and if $\boldsymbol{\Xi}$ is an analytic function on U taking values in V and establishing between U and V a one-to-one nonsingular correspondence with analytic inverse, then we shall say that $\boldsymbol{\Xi}$ is an analytic correspondence between U and V.

Observation. Some among $d, d', \overline{d}, \overline{d}'$, may vanish: see the Observation to Definition 14.

It is now possible to establish the definition of an analytic surface. The reader should try to make drawings to see the various geometrical objects discussed in the following definitions and observations.

16 Definition. *A regular surface $\Sigma \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ is said to be "locally analytic" in an open set $U \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ if there is a family of local analytic systems of coordinates $(U_\alpha, \boldsymbol{\Xi}_\alpha)_{\alpha \in A}$ with bases $(\Omega_\alpha)_{\alpha \in A}$ such that:*
(i) the points of $\Sigma \cap U_\alpha$ are those which in $(U_\alpha, \boldsymbol{\Xi}_\alpha)$ have coordinates $\beta_1 = \ldots = \beta_{d+d'-\ell} = 0$ where ℓ is the "dimension" of Σ, i.e., $(U_\alpha, \boldsymbol{\Xi}_\alpha)$ are adapted to Σ;
(ii) as α varies in A, the sets U_α cover $\Sigma \cap U$ and A is a finite set of indices.
If Σ is a locally analytic surface in U and f is an $\mathcal{R}^d \times \mathcal{T}^{d'}$-valued function on Σ, we shall say that "f is analytic on Σ" if it is the restriction to Σ of an analytic function on an open set $\widetilde{U} \supset \Sigma \cap U$.
If $\Sigma \subset U$ is a closed set and if Σ is a locally analytic surface in U, we shall say that Σ is an "analytic surface" (this notion is U independent).

Observations.
(1) If some of the $d, d', \overline{d}, \overline{d}'$ vanish see the Observation to Definition 14.
(2) Examples are discussed in the problems and exercises at the end of this section.

Finally, we define the "analytically regular sets".

17 Definition. *A closed set $G \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ will be called "locally analytic" in the open set $U \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ if ∂G is a surface locally analytic in U.*
If G is locally analytic in U and $G \subset U$, then G will be called "an analytic set" (this notion is U independent).
A closed set $G \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ will be called "analytically regular" if there is an open set $U \supset G$ and a family of sets locally analytic in U through which, via a finite number of union and intersection operations, one can build G.

Observations.
(1) If d or d' vanish, see comment (1) to Definition 14.
(2) Any analytic surface is an analytic set (since either $\partial\Sigma = \Sigma$ or $\Sigma \equiv \mathcal{R}^d \times \mathcal{T}^{d'}$.
(3) If $\boldsymbol{\Xi}$ is an analytic transformation of $U \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ onto $V \subset \mathcal{R}^d X \mathcal{T}^{d'}$ and if $G \subset U$ is an analytically regular set then $\boldsymbol{\Xi}(G) \subset V$ is also analytically regular, i.e., the above notion is invariant under analytic maps. This follows from the fact that composing analytic functions, one obtains analytic func-

tions.[16]
(4) If Σ is a surface locally analytic in U and if $G \subset U$ is an analytically regular set also, $G \cap \Sigma$ is an analytically regular set: this is the "invariance under the intersection operations" of the analytic regularity.
(5) Let $\overline{d} \geq d, \overline{d'} \geq d'$ and regard $\mathcal{R}^d \times \mathcal{T}^{d'}$ as a subset of $\mathcal{R}^{\overline{d}} \times \mathcal{T}^{\overline{d'}}$ by identifying it as the subset of $\mathcal{R}^{\overline{d}} \times \mathcal{T}^{\overline{d'}}$ consisting of the points $(\overline{\mathbf{x}}\overline{\boldsymbol{\varphi}})$ such that $\overline{\mathbf{x}} = (\mathbf{x}, \mathbf{0})$, $\overline{\boldsymbol{\varphi}} = (\boldsymbol{\varphi}, \mathbf{0})$ with $(\mathbf{x}, \boldsymbol{\varphi}) \in \mathcal{R}^d \times \mathcal{T}^{d'}$ and the $\mathbf{0}$'s denote the origins in $\mathcal{R}^{\overline{d}-d}$ and $\mathcal{R}^{\overline{d'}-d}$, respectively. Then $\mathcal{R}^d \times \mathcal{T}^{d'}$ is an analytic surface in $\mathcal{R}^{\overline{d}} \times \mathcal{T}^{\overline{d'}}$
If $G \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ is analytically regular, then its "extension $\widehat{G}$ to $\mathcal{R}^{\overline{d}} \times \mathcal{T}^{\overline{d}'}$, $\widehat{G} = \{(\mathbf{x}, \boldsymbol{\varphi}) \,|\, (\overline{\mathbf{x}}, \overline{\boldsymbol{\varphi}}) \in \mathcal{R}^{\overline{d}} \times \mathcal{T}^{\overline{d}'}, \overline{\mathbf{x}} = (\mathbf{x}, \mathbf{y}), \ \overline{\boldsymbol{\varphi}} = (\boldsymbol{\varphi}, \boldsymbol{\psi}) \text{ with } (\mathbf{x}, \boldsymbol{\varphi}) \in G\}$ is analytically regular in $\mathcal{R}^d \times \mathcal{T}^{d'}$.
(6) On every regular (or locally analytic) surface $\Sigma \subset \mathcal{R}^d \times \mathcal{T}^{d'}$, one can define the "area measure": if $(U, \boldsymbol{\Xi})$ is a regular (or analytic) system of local coordinates adapted to Σ with basis Ω, there is a regular (or analytic) function σ on Ω such that for $E \subset\in \Sigma \cap U$:

$$\text{area}(E) = \int_{\boldsymbol{\Xi}^{-1}(E)} \sigma(0, \ldots, 0, \beta_{d+d'-\ell+1}, \ldots, \beta_{d+d'}) d\beta_{d+d'-\ell+1} \ldots d\beta_{d+d'} \tag{4.13.12}$$

provided $\boldsymbol{\Xi}^{-1}(E)$ is measurable in the Riemannian sense (in this case, one says that E is measurable with respect to the area measure).
If $\Sigma \subset \mathcal{R}^d$ is a regular surface and $(U, \boldsymbol{\Xi})$ is a well-adapted orthogonal and of Fermi type system of local coordinates (in the sense of Definition 12, §3.7, p.177, and Proposition 12, p.183) with respect to the scalar product $\boldsymbol{\eta} \cdot \boldsymbol{\chi}$ on $\mathcal{R}^d$ one has

$$\sigma(0, \ldots, 0, \beta_{d+d'-\ell+1}, \ldots, \beta_{d+d'}) = \sqrt{\gamma^\ell} \tag{4.13.13}$$

essentially by (a very reasonable) definition; σ in the other coordinate systems is computed by ordinary coordinate transformations. The simplicity of Eq. (4.13.13) provides a further illustration of the notion of "well-adapted orthogonal" systems of coordinates of §3.7.
(7) One may think that it is possible to define something like "C^∞-regular" sets by simply replacing the word analytic by C^∞ everywhere above. However, the property in Observation (4), for instance, would not hold. See exercises to §4.13.

The problem of the (Riemann) measurability of sets is not always trivial and the interest in the above digression on the definition of analytically regular sets rests mainly on the validity of the following proposition.

28 Proposition. *Let $\Sigma \subset \mathcal{R}^d \times \mathcal{T}^{d'}$ be a surface locally analytic in the open set U and let E be the analytically regular set contained in U.*

[16] The reader can attempt a proof starting with the $\ell = 1$ case.

The set $E \cap \Sigma$ is then measurable with respect to the area measure on Σ, i.e., given $\varepsilon > 0$, there exist two functions χ^+ and χ^- of class C^∞ on Σ, $0 \le \chi^- \le \chi^+ \le 1$, such that if χ_E, is the characteristic function of $E \cap \Sigma$:

$$\text{(i)} \qquad \chi^-(\boldsymbol{\xi}) \le \chi_E(\boldsymbol{\xi}) \le \chi^+(\boldsymbol{\xi}), \forall\, \boldsymbol{\xi} \in \Sigma \cap E, \tag{4.13.14}$$

$$\text{(ii)} \qquad \int_{E\cap\Sigma} (\chi^+(\boldsymbol{\xi}) - \chi^-(\boldsymbol{\xi}))\, d\sigma_{\boldsymbol{\xi}} < \varepsilon, \tag{4.13.15}$$

where the integral denotes the surface integral on Σ, see Observation (6) above.

We do not describe the proof of this proposition. Although it is not particularly difficult it would require a preliminary analysis of the structure of the analytic surfaces and their intersections which, being marginal for us, would lead us too far away from our problem of discussing the integrability criteria for mechanical systems.

addcontentslinetocsubsectionExercises and Problems

1. Show that the function on $\mathcal{T}^1$: $\varphi \to \cos\varphi$ is analytic.

2. Show that the $\mathcal{T}^1$-valued function $x \to x \bmod 2\pi$ is analytic.

3. If $f \in C^\infty(T^1)$ and if its Fourier coefficients can be bounded as $|\widehat{f}_k| \le Fc^k$, $c < 1$, then f is analytic on V. Prove this statement.

4. Generalize Problem 3 to the case of a function on $\mathcal{T}^\ell$.

5. Show that a "surface" of $\mathcal{R}$ relatively closed in $U \subset \mathcal{R}$ and locally analytic in U (U open) is, inside U, a union of at most denumerably many points without accumulation points in U or coincides with U. Show that a bounded analytically regular set in $\mathcal{R}$ is a union of finitely many points and closed intervals.

6. Show that straight lines, planes, half-lines, half-planes, and half-spaces are analytic sets in $\mathcal{R}^2$ and $\mathcal{R}^3$.

7. Show that triangles, polygons, disks and their boundaries are analytically regular in $\mathcal{R}^2$ and in $\mathcal{R}^3$.

8. Show that the regular solids, the spheres, the diedra, the triedra, etc., and their boundaries are analytically regular in $\mathcal{R}^3$.

9. Show that the disk and the ellipse, or the ball and the ellipsoidal ball (i.e. the sets whose boundary are the sphere or the ellipsoid), and their boundaries are analytic sets in $\mathcal{R}^2$ or $\mathcal{R}^3$, respectively.

10. Show that a disk in $\mathcal{R}^3$ is not an analytic set although it is analytically regular.

11. Let $x_1, x_2, \ldots$ be a numeration of the rational numbers in $[0, 1]$. For every x_k, consider the open interval with length 2^{-1-k} and center x_k. Show that the union of such intervals is an open set dense in $[0, 1]$ with external measure (in the Riemannian sense) ≥ 1 and with the internal measure $< \frac{1}{2}$. Call this union A.

12. Let $g \in C^\infty(\mathcal{R})$ be a positive function on $(-\frac{1}{2}, \frac{1}{2})$ and zero elsewhere. Set

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k!} g(2^{k+1}(x - x_k))$$

and show that f is positive on A (see Problem 11) and zero outside. Show also that f is in $C^\infty(\mathcal{R})$.

13. Show that the set $D_\infty \stackrel{def}{=} \{(x,y)\,|\,0 \le x \le 1,\ f(x) \le y < +\infty \subset \mathcal{R}^2$, with f as defined in Problem 12, has a piecewise C^∞ boundary. Show that the intersection between D_∞ and the x-axis is not measurable in the Riemannian sense.

14. Let $v \in C^\infty(\mathcal{R}^1)$ and consider the surface in $\mathcal{R}^2$ with equations $z = v(x)$, $x \in \mathcal{R}$. Show that the "area" of a line element over dx (i.e. its length) is, according to Eq. (4.13.13), $d\sigma = \sqrt{1 + (\frac{dv(x)}{dx})^2}dx$.

15. Let $v \in C^\infty(\mathcal{R}^2)$ and consider the surface in $\mathcal{R}^3$ with equations $z = v(x,y)$, $(x,y) \in \mathcal{R}^2$. Show that the area of a surface element over $dxdy$ is, according to Eq. (4.13.13),

$$d\sigma = \sqrt{1 + (\frac{\partial v(x,y)}{\partial x})^2 + (\frac{\partial v(x,y)}{\partial y})^2}dxdy.$$

4.14 Analytically Integrable Systems. Frequency of Visits and Ergodicity

In §4.8, Definition 11, p.289, we introduced the notion of "analytically integrable" Hamiltonian systems defined on an open set $W \subset \mathcal{R}^\ell \times \mathcal{R}^\ell$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times \mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2}, \ell_1 + \ell_2 = \ell$, in phase space.

The interest in analytically integrable systems is twofold: essentially all concrete integrable systems so far met were analytically integrable (and this could be verified with some labor); furthermore, if $\mathbf{t} = (it_1)_{i=0}^\infty$, the $(f,\mathbf{t})$ histories of the points in the integrability region W of phase space have a well-defined average value for all the $f \in C^\infty(W)$, and also for many other more singular functions f, for instance for the characteristic functions of the analytically regular sets.

To illustrate this remarkable property, it is convenient to introduce the following notions.

18 Definition. *Let W be a subset of the phase space ($\subset \mathcal{R}^\ell \times \mathcal{R}^\ell$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times \mathcal{R}^{\ell_1}\mathcal{T}^{\ell_2}, \ell_1 + \ell_2 = \ell$) of an analytic time-independent Hamiltonian system. Suppose that the system is analytically integrable on W.*
Let $\mathcal{E} = (E_0, E_1 1, \ldots, E_p)$ be a family of subsets of W such that:
(i) $\cup_{i=0}^p E_i = W$, $E_i \cap E_j = \emptyset$ if $i \ne j$, i.e., $\mathcal{E}$ is a "partition of W";
(ii) $E_1, \ldots, E_p$ are analytically regular;
(iii) $d(E_i, E_j) > 0$ if $i \ne j$, $i, j = 1, \ldots, p$.
Obviously, $E_0 = W \backslash \cup_{j=1}^p E_j$ is an open set.
The partition $\mathcal{E}$ will be called an "analytically regular partition" of W. Denote χ_{E_i} the characteristic function of the sets $E_i, i =, 0, 1, \ldots, p$, and let

$$f_\mathcal{E}(\boldsymbol{\xi}) = \sum_{j=0}^p j\,\chi_{E_j}(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in W \tag{4.14.1}$$

and, finally, we call "$(G,\mathbf{t})$ history of $(\mathbf{p},\mathbf{q}) \in W$" the $(f_\mathcal{E},\mathbf{t})$ history of $(\mathbf{p},\mathbf{q})$ when $t = (t_i)_{i=0}^\infty$ is a divergent monotonic sequence.

Observation. The $(\mathcal{E}, \mathbf{t})$ history is a sequence of integers between 0 and p: the k-th element of this sequence simply indicates into which set among those of $\mathcal{E}$ the point $S_{t_k}(\boldsymbol{\xi})$ falls, if $t \to S_t(\boldsymbol{\xi})$, $t \geq 0$, is the solution to the Hamiltonian equations with initial datum $\boldsymbol{\xi} =_(\mathbf{p}, \mathbf{q})$.

The following proposition is very remarkable.

29 Proposition. *Given an analytic Hamiltonian system analytically integrable in the subset W of phase space, the limit*

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{j=0}^{N-1} \chi_E(S_{jt_1}(\boldsymbol{\xi})) \tag{4.14.2}$$

exists, $\forall\, t_1 > 0$, $\forall\, \boldsymbol{\xi} \in W$ and for all analytically regular subsets E of W. This limit will be called naturally the "frequency of visit to E by the motion starting in $\boldsymbol{\xi}$" with respect to the sequence of observation times $\mathbf{t} = (it_1)_{i=0}^{\infty}$, $t_1 > 0$.

Proof. The image $I(E) \subset V \times \mathcal{T}^\ell$ of E via the analytic integrating transformation I, see Definition 11, §4.8, p.289, will still be analytically regular, see observation (3) to Definition 17, p.338. Since for $I(\boldsymbol{\xi}) = (\mathbf{A}, \boldsymbol{\varphi})$,

$$I(S_t\boldsymbol{\xi}) = (\mathbf{A}, \boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{A})t), \tag{4.14.3}$$

the proof of the above proposition is "reduced" to the one contemplated in the following one.

30 Proposition. *Let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_\ell)$ be an ℓ-tuple of real numbers and let $E \subset \mathcal{T}^\ell$ be an analytically regular subset of $\mathcal{T}^\ell$. If $\mathbf{t} = (it_1)_{i=0}^{\infty}$, $t_1 > 0$, the frequency of visits*

$$\nu_E(\boldsymbol{\varphi}) = \lim_{N\to+\infty} \frac{1}{N} \sum_{j=0}^{N-1} \chi_E(\boldsymbol{\varphi} + \boldsymbol{\omega}t) \tag{4.14.4}$$

exists, $\forall\, \boldsymbol{\varphi} \in \mathcal{T}^\ell$. Furthermore, if the numbers $\omega_1, \ldots, \omega_\ell$ and $\sigma = \frac{2\pi}{t_1}$ are rationally independent,

$$\nu_E(\boldsymbol{\varphi}) = \frac{1}{(2\pi)^\ell} \int_{\mathcal{T}^\ell} d\boldsymbol{\varphi}' \tag{4.14.5}$$

Proof. We shall only treat the simple case when $(\omega_1, \ldots, \omega_\ell, \sigma)$ are rationally independent because it is easy. The general case can be reduced to this one with some patient though interesting work which we leave to the reader, referring, as a guide, to the sequence of problems at the end of this section.

The idea of the proof is to use the Riemann measurability of E (consequence of its analytic regularity, see Proposition 28) to find two $C^\infty(\mathcal{T}^\ell)$ functions χ^- and χ^+ verifying Eqs. (4.13.14) and (4.13.15) to infer that $\nu_E(\boldsymbol{\varphi})$, if existing, must be between the averages of χ^- and χ^+ which, in turn, exist and

differ at most by ε by Eq. (4.13.15) and Proposition 6, §4.2, Eq. (4.2.10), p.251. Then the arbitrariness of ε implies the actual existence of $\nu_E(\boldsymbol{\varphi})$ and the fact that it is between the averages $\frac{1}{(2\pi)^\ell}\int_{\mathcal{T}^\ell}\chi^{\pm}(\boldsymbol{\varphi}')\,d\boldsymbol{\varphi}'$. Again the arbitrariness of ε and Eqs. (4.13.15) and (4.13.14) imply Eq. (4.14.5). mbe

Observations.
(1) The reader will note that the analytic regularity of E is used in the above proof only to infer the Riemann measurability of E. However, if $\omega_1,\ldots,\omega_\ell,\sigma$ were not rationally independent the analytic regularity should again be used to prove the reducibility of the general case to the rationally independent one. This is the reason why the Riemann measurability is not in itself a genera] sufficient condition for the existence of the limit of Eq. (4.14.4). See problems at the end of this section.
(2) Of course if $\omega_1,\ldots,\omega_\ell,\sigma$ are rationally independent the Riemann measurability of E suffices, alone, to deduce Eqs. (4.14.5) and (4.14.6) as it appears clear from the above proof.

So every motion of a Hamiltonian system analytically integrable in W visits an analytically regular set E with a well-defined frequency of visit. One can wonder about the frequency of joint visits to two given analytically regular sets E and E'. The remarkable fact is that they are, on the average, "independent". The frequency of a visit to E followed j time units later by a visit to E' is, on the average over j, equal to the product of the frequency of visit to E and of that of E': $\forall\,\boldsymbol{\xi}\in W$, $\forall\,t_1>0$,

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{j=0}^{N-1}\nu_{E\cap S_{jt_1}(E')}(\boldsymbol{\xi})=\nu_E(\boldsymbol{\xi})\nu_{E'}(\boldsymbol{\xi}). \tag{4.14.6}$$

In other words, visit to E by a given motion does not put any restrictions on the possibility of a visit to E' j time units later, at least on the average on j.

This is the content of the following proposition.

31 Proposition. *In the assumptions of Proposition 29, let $E,E'\subset W$ be two analytically regular sets. Then property (4.14.6) holds for all $\boldsymbol{\xi}\in W$, $\forall\,t_1>0$.*

Observation. This proposition is a corollary of the following Proposition 32 on the quasi-periodic motions on $\mathcal{T}^\ell$ in the same way in which Proposition 29 appears to be a corollary of Proposition 30.

32 Proposition. *Let $E,E'\subset\mathcal{T}^\ell$ be two analytically regular sets and let $\boldsymbol{\omega}\in\mathcal{R}^\ell$, $t_1>0$. Denote $E'+t\boldsymbol{\omega}$ the set of points $\boldsymbol{\varphi}'+t\boldsymbol{\omega}\mod 2\pi$ as $\boldsymbol{\varphi}'$ varies in E' ($E'+t\boldsymbol{\omega}$ is the set into which E' evolves in time t under the quasi periodic flow on $\mathcal{T}^\ell$ with pulsations $\boldsymbol{\omega}$). If $\nu_E(\boldsymbol{\varphi})$ is the frequency of visits of the points $\boldsymbol{\varphi}+jt_1\boldsymbol{\omega}$, $j=0,1,\ldots$ to E, it is, $\forall\,\boldsymbol{\varphi}\in\mathcal{T}^\ell$,*

$$\lim_{N\to+\infty}\frac{1}{N}\sum_{j=0}^{N-1}\nu_{E\cap(E'+jt_1\boldsymbol{\omega})}(\boldsymbol{\varphi})=\nu_E(\boldsymbol{\varphi})\nu_{E'}(\boldsymbol{\varphi}). \tag{4.14.7}$$

Observations
(1) When $\omega_1, \ldots, \omega_\ell$ and $\sigma = \frac{2\pi}{t_1}$ are rationally independent $\nu_E(\boldsymbol{\varphi})$ is the measure of E, see Eq. (4.14.5). Hence, Eq. (4.14.7) means that in this case the fraction of E occupied by images of points of E' is a fraction of E equal, on the average, to the measure of E'. In other words, $E' + jt_1\boldsymbol{\omega}$ is uniformly scattered in $\mathcal{T}^\ell$, on the average. This holds for $\forall\, E'$ analytically regular.
(2) By considering the case $\ell = 1$ and taking E and E' to be two small intervals, one sees that the limit of $\nu_{E\cap(E'+jt_1\boldsymbol{\omega})}(\boldsymbol{\varphi})$ as $j \to \infty$ does not exist in general: even in the case of rational independence of $\boldsymbol{\omega}$, $\frac{2\pi}{t_1}$ the average over j in Eq. (4.14.7) is essential. Therefore, even though on the average $E' + jt_1\boldsymbol{\omega}$ is uniformly scattered in $\mathcal{T}^\ell$, it is not true that for large times j this set is uniformly scattered. This is due manifestly to the fact that the rotations of the torus are "rigid" transformations and they do not "mix" the points of $\mathcal{T}^\ell$.

PROOF. As in the case of Proposition 29, let us only treat the simple case when $\omega_1, \ldots, \omega_\ell$ and $\sigma = \frac{2\pi}{t_1}$ are rationally independent. The general case can be treated by solving the last of the problems at the end of this section.

Proceeding as in the proof of Proposition 30 and using the Riemann measurability [see Eq. (4.13.15)] of the sets E, E', the problem of proving Eq. (4.14.7) is reduced to that of proving, $\forall\, f, g \in C^\infty(\mathcal{T}^\ell)$,

$$\lim_{N\to+\infty} \frac{1}{N} \sum_{j=0}^{N-1} \overline{f(\boldsymbol{\varphi})g(\boldsymbol{\varphi} + jt_1\boldsymbol{\omega})} = \overline{f(\boldsymbol{\varphi})}\,\overline{g(\boldsymbol{\varphi})} \tag{4.14.8}$$

where the bar over a function of $\boldsymbol{\varphi}$ denotes the average:

$$\overline{f(\boldsymbol{\varphi})} = \lim_{N\to+\infty} \frac{1}{N} \sum_{j=0}^{N-1} f(\boldsymbol{\varphi} + jt_1\boldsymbol{\omega}). \tag{4.14.9}$$

Note that Eq. (4.14.8) would directly become Eq. (4.14.7) if one could take $f = \chi_E$, $g = \eta_{E'}$.

To prove this proposition, Eq. (4.14.8) shall be applied to the functions χ^+, χ'^+ which, according to Proposition 28, approximate $\chi_E, \chi_{E'}$ from above and to the functions χ^-, χ'^- which approximate $\chi_E, \chi_{E'}$ from below, following the approximation idea of the proof of Proposition 30.

Eq. (4.14.8) can now be checked. By the simplifying assumption of rational independence, see Proposition 6, §4.2, p.251,

$$\overline{f(\boldsymbol{\varphi})} = \int_{\mathcal{T}^\ell} f(\boldsymbol{\varphi}') \frac{d\boldsymbol{\varphi}'}{(2\pi)^\ell} = \widehat{f}_{\mathbf{0}}, \qquad \overline{g(\boldsymbol{\varphi})} = \int_{\mathcal{T}^\ell} g(\boldsymbol{\varphi}') \frac{d\boldsymbol{\varphi}'}{(2\pi)^\ell} = \widehat{g}_{\mathbf{0}}, \tag{4.14.10}$$

$\widehat{f}_{\mathbf{n}}, \widehat{g}_{\mathbf{n}}$ being the Fourier coefficients of f, g, respectively. Furthermore, since

$$
\begin{aligned}
f(\boldsymbol{\varphi})g(\boldsymbol{\varphi}+jt_1\boldsymbol{\omega}) &= \sum_{\mathbf{n},\mathbf{n}'} \widehat{f}_{\mathbf{n}}\,\widehat{g}_{\mathbf{n}'} e^{i(\mathbf{n}\cdot\boldsymbol{\varphi}+\mathbf{n}'\cdot\boldsymbol{\varphi}+jt_1\mathbf{n}'\cdot\boldsymbol{\omega})} \\
&= \sum_{\mathbf{m}} e^{i\mathbf{m}\cdot\boldsymbol{\varphi}}\Big(\sum_{\mathbf{n}+\mathbf{n}'=\mathbf{m}} \widehat{f}_{\mathbf{n}}\,\widehat{g}_{\mathbf{n}'} e^{i\,jt_1\mathbf{n}'\cdot\boldsymbol{\omega}}\Big)
\end{aligned}
\tag{4.14.11}
$$

one finds, still by Proposition 6, §4.2,

$$
\begin{aligned}
\overline{f(\boldsymbol{\varphi})g(\boldsymbol{\varphi}+jt_1\boldsymbol{\omega})} =& \text{Fourier coefficient of order } \mathbf{0} \\
& \text{of the function in Eqs. (4.14.11)} \\
=& \sum_{\mathbf{n}} \widehat{f}_{\mathbf{n}}\,\widehat{g}_{-\mathbf{n}} e^{-i\,jt_1\mathbf{n}\cdot\boldsymbol{\omega}}
\end{aligned}
\tag{4.14.12}
$$

Then $\frac{1}{N}\sum_{j=0}^{N-1}\overline{f(\boldsymbol{\varphi})g(\boldsymbol{\varphi}+jt_1\boldsymbol{\omega})}$ is

$$
\sum_{\mathbf{n}} \widehat{f}_{\mathbf{n}}\,\widehat{g}_{-\mathbf{n}} \frac{1}{N}\sum_{j=0}^{N-1} e^{-i\,jt_1\mathbf{n}\cdot\boldsymbol{\omega}} = \widehat{f}_{\mathbf{0}}\widehat{g}_{\mathbf{0}} + \sum_{\mathbf{n}\neq\mathbf{0}} \widehat{f}_{\mathbf{n}}\,\widehat{g}_{-\mathbf{n}} \frac{1}{N}\frac{e^{-i\,Nt_1\mathbf{n}\cdot\boldsymbol{\omega}}-1}{e^{-i\,t_1\mathbf{n}\cdot\boldsymbol{\omega}}-1} \tag{4.14.13}
$$

and, by the usual argument of passage to the limit under the series sign, it follows that the limit as $N \to +\infty$ of Eq. (4.14.13) is just $\widehat{f}_{\mathbf{0}}\widehat{g}_{\mathbf{0}}$ which shows, recalling (4.14.10), the validity of Eq. (4.14.8) and, hence, the above proposition validity (in the special case treated here). mbe

The above propositions imply some simple consequences.
Let $\mathcal{E} = (E_0, E_1, \ldots, E_s)$ be a partition of the phase space W of an analytically integrable Hamiltonian system. Suppose that $\mathcal{E}$ is analytically regular in W in the sense of Definition 18, p.341. Given $\mathbf{t} = (i\,t_1)_{i=0}^{\infty}$, the partition $\mathcal{E}$ and $k \geq 0$, $0 \leq j_1 < j_2 < \ldots < j_k$, define

$$
E\begin{pmatrix} j_1 & \ldots & j_k \\ \alpha_1 & \ldots & \alpha_k \end{pmatrix} \stackrel{def}{=} S_{-j_1t_1}(E_{\alpha_1}) \cap S_{-j_2t_1}(E_{\alpha_2}) \cap \ldots \cap S_{-j_kt_1}(E_{\alpha_k}). \tag{4.14.14}
$$

This is the set of the points $\boldsymbol{\xi} \in W$ such that

$$
S_{j_1t_1}(\boldsymbol{\xi}) \in E_{\alpha_1},\, S_{j_2t_1}(\boldsymbol{\xi}) \in E_{\alpha_2}, \ldots, S_{j_kt_1}(\boldsymbol{\xi}) \in E_{\alpha_k}. \tag{4.14.15}
$$

From Eq. (4.14.15) and from the fact that $\mathcal{E}$ is a partition of W, it is

$$
E\begin{pmatrix} j_1 & \ldots & j_k \\ \alpha_1 & \ldots & \alpha_k \end{pmatrix} \cap E\begin{pmatrix} j_1 & \ldots & j_k \\ \beta_1 & \ldots & \beta_k \end{pmatrix} = \emptyset \tag{4.14.16}
$$

unless $\alpha_1 = \beta_1, \ldots, \alpha_k = \beta_k$. Also

$$
\bigcup_{\alpha_1,\ldots,\alpha_k}^{0,s} E\begin{pmatrix} j_1 & \ldots & j_k \\ \alpha_1 & \ldots & \alpha_k \end{pmatrix} = W \qquad \text{and} \tag{4.14.17}
$$

$$\bigcup_{\alpha}^{0,s} E\begin{pmatrix} j_1 \dots j_{p-1}\, j_p\, j_{p+1} \, \dots \, j_k \\ \alpha_1 \dots \alpha_{p-1}\, \alpha\, \alpha_{p+1} \, \dots \, \alpha_k \end{pmatrix} = E\begin{pmatrix} j_1 \dots j_{p-1}\, j_{p+1} \, \dots \, j_k \\ \alpha_1 \dots \alpha_{p-1}\, \alpha_{p+1} \, \dots \, \alpha_k \end{pmatrix} \quad (4.14.18)$$

It is also clear that if $\alpha_i \neq 0, \forall\, i =, \dots, k$, the set $E\begin{pmatrix} j_1 & \dots & j_k \\ \alpha_1 & \dots & \alpha_k \end{pmatrix}$ is analytically regular because the time evolution transformations $(S_t)_{t\in\mathcal{R}}$ are analytic (being such after the analytic change of coordinates I, [see Eq. (4.8.14)] which integrates the system,[17] and because the analytic image of an analytically regular set is still analytically regular [see Observation (3) to Definition 17, p.338).

The sets $E\begin{pmatrix} j_1 & \dots & j_k \\ \alpha_1 & \dots & \alpha_k \end{pmatrix}$ have a simple physical meaning.

We imagine that the partition $\mathcal{E}$ models an actual observation of some physical quantity. The results of the observations, read on a dial, give a finite number of results $1, 2, \dots, s$ or 0 ("off the dial"). Since the results of physical measurements can always be numbered from 1 to some s, this is a very general model.

Thus the phase space W is divided by collecting together all the physical configurations $\boldsymbol{\xi} \in W$ that produce the same result for the value of the physical quantity described by Eq. (4.14.1) in this model.

Given a sequence of observation times $0, t_1, t_2, \dots, t_j = jt_1$, we can decide to record the results of the observations made at times $j_1 t_1, \dots, j_k t_1$. We see that the possible outcomes of such observations are $(s+1)^k$ k-tuples $(\alpha_1, \dots, a_k)$ and we can partition W into $(s+1)^k$ sets of the form $E\begin{pmatrix} j_1 & \dots & j_k \\ \alpha_1 & \dots & \alpha_k \end{pmatrix}$ collecting the points falling in E_{α_1} at time $j_1 t_1$, in E_{α_2} at time $j_2 t_1, \dots$, in E_{α_k} at time $j_k t_1$.

In terms of the above mathematical notions, it is possible to formulate an interesting proposition whose physical meaning can easily be gathered from the just discussed interpretation.

33 Proposition. *Let W be the phase space of a time-independent analytically integrable Hamiltonian system. Let $t_1 > 0$ and let $\mathcal{E} = (E_0, \dots, E_s)$ be an analytically regular partition of W. Then $\forall\, \boldsymbol{\xi} \in W$, the frequencies of visits to the sets of the form of Eq. (4.14.14) by the motion starting at $\boldsymbol{\xi}$ exist. Denoting such frequencies*

$$\mathbf{p}\begin{pmatrix} j_1 \dots j_k \\ \alpha_1 \dots \alpha_k \end{pmatrix}\!\Big|\boldsymbol{\xi} \Big) \stackrel{def}{=} \nu_{\left(\begin{smallmatrix} j_1 \dots j_k \\ \alpha_1 \dots \alpha_k \end{smallmatrix}\right)}(\boldsymbol{\xi}) \quad (4.14.19)$$

it also follows that:
(i) $\forall\, k \in \mathcal{Z}_+, \forall\, 0 \le j_1 < j_2 < \dots < j_k$ *integers,* $\forall\, \alpha_1, \alpha_2, \dots, \alpha_k$ *in* $(0, 1, \dots, s)$

[17] Here we use a well-known fact that when composing analytic functions, one obtains analytic functions. The reader can attempt a proof of this starting with the $\ell = 1$ case.

$$\mathbf{p}\begin{pmatrix} j_1 \cdots j_k \\ \alpha_1 \ldots \alpha_k \end{pmatrix}\!\!\left|\boldsymbol{\xi}\right) \stackrel{def}{=} \nu_{\binom{j_1 \cdots j_k}{\alpha_1 \cdots \alpha_k}}(\boldsymbol{\xi}) \geq 0 \tag{4.14.20}$$

(ii) $\forall\, k \in \mathcal{Z}_+, \forall\, 0 \leq j_1 < j_2 < \ldots < j_k$ *integers,* $\forall\, \alpha_1, \alpha_2, \ldots, \alpha_k$ *in* $(0, 1, \ldots, s)$*,* $\forall\, p = 1, 2, \ldots, k$*:*

$$\sum_{\alpha=0}^{s} \mathbf{p}\left(\begin{matrix} j_1 \ldots j_{p-1}\, j_p\, j_{p+1} \ \cdots\ j_k \\ \alpha_1 \ldots \alpha_{p-1}\, \alpha\, \alpha_{p+1} \ \ldots\ \alpha_k \end{matrix}\middle|\boldsymbol{\xi}\right) = \mathbf{p}\left(\begin{matrix} j_1 \cdots j_{p-1}\, j_{p+1} \ \cdots\ j_k \\ \alpha_1 \ldots \alpha_{p-1}\, \alpha_{p+1} \ \ldots\ \alpha_k \end{matrix}\middle|\boldsymbol{\xi}\right) \tag{4.14.21}$$

(iii) $\forall\, k \in \mathcal{Z}_+, \forall\, 0 \leq j_1 < j_2 < \ldots < j_k$ *integers:*

$$\sum_{\alpha_1, \ldots, \alpha_k} \mathbf{p}\left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix}\middle|\boldsymbol{\xi}\right) \equiv 1 \tag{4.14.22}$$

(iv) $\forall\, k \in \mathcal{Z}_+, \forall\, 0 \leq j_1 < j_2 < \ldots < j_k$*,* $0 \leq i_1 < i_2 < \ldots < i_h$ *integers, and* $\forall\, \alpha_1, \alpha_2, \ldots, \alpha_k$*,* $\beta_1, \beta_2, \ldots, \beta_h$ *in* $(0, 1, \ldots, s)$

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \sum_{\ell=0}^{N-1} \mathbf{p}\left(\begin{matrix} j_1 \ldots j_k\, i_{1+\ell} \ldots i_{h+\ell} \\ \alpha_1 \ldots \alpha_k\, \beta_1 \quad \ldots \beta_h \end{matrix}\middle|\boldsymbol{\xi}\right) \\ &= \mathbf{p}\left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix}\middle|\boldsymbol{\xi}\right) \mathbf{p}\left(\begin{matrix} i_1 \ldots i_h \\ \beta_1 \ldots \beta_h \end{matrix}\middle|\boldsymbol{\xi}\right). \end{aligned} \tag{4.14.23}$$

Properties (i), (ii), (iii), and (iv) will be referred to, respectively, as "positivity", "compatibility", "normalization", and "ergodicity" properties of the frequencies of the motion "generated by $\boldsymbol{\xi}$ *and observed on* $\mathcal{E}$*".*

Observation. It will appear that the above proposition is just a fancy statement of the results already obtained. However, it is very useful because it introduces a few qualitative notions which are very natural and important.

Proof. First suppose the existence of the frequencies of Eq. (4.14.19). Then (i) is obvious, while (ii) and (iii) follow from Eqs. (4.14.16)-(4.14.18) and Eqs. (4.14.16) and (4.14.17), respectively.

So it remains to prove the existence of the frequencies and (iv). The existence of the frequencies for the sets $E\binom{j_1 \ldots j_k}{\alpha_1 \ldots \alpha_k}$ with $\alpha_1 \neq 0, \ldots, \alpha_k \neq 0$ follows from their analytic regularity stated after their definition and from Proposition 29. It remains, therefore, to examine the cases when some among $\alpha_1, \ldots, \alpha_k$ are 0.

Proceeding inductively from Eqs. (4.14.16), (4.14.18), note that if $k = 1$, Eq. (4.14.17) and the definition of frequency imply existence of $\mathbf{p}\left(\begin{smallmatrix} j_1 \\ 0 \end{smallmatrix}\middle|\boldsymbol{\xi}\right)$ and, actually,

$$\mathbf{p}\left(\begin{matrix} j_1 \\ 0 \end{matrix}\middle|\boldsymbol{\xi}\right) = 1 - \sum_{\alpha=1}^{s} \mathbf{p}\left(\begin{matrix} j_1 \\ \alpha \end{matrix}\middle|\boldsymbol{\xi}\right). \tag{4.14.24}$$

In fact, in general, if E is visited with well defined frequency ν its complement is visited with frequency equal to $(1-\nu)$. If $k=2$, by the same arguments, we deduce that for $a>0$,

$$\mathbf{p}\begin{pmatrix} j_1\, j_2 \\ 0\,\alpha \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big) = \mathbf{p}\begin{pmatrix} j_2 \\ \alpha \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big) - \sum_{\alpha'=1}^{s} \mathbf{p}\begin{pmatrix} j_1\, j_2 \\ \alpha'\,\alpha \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big). \tag{4.14.25}$$

Hence the frequency

$$\mathbf{p}\begin{pmatrix} j_1\, j_2 \\ 0\,0 \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big) = \mathbf{p}\begin{pmatrix} j_1 \\ 0 \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big) - \sum_{\alpha=1}^{s} \mathbf{p}\begin{pmatrix} j_1\, j_2 \\ 0\,\alpha \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big). \tag{4.14.26}$$

exists for the same reasons, etc., inductively.

Finally, (iv) follows, by Proposition 31, immediately when $\alpha_1 \neq 0, \ldots, \alpha_k \neq 0, \beta_1, \ldots, \beta_h \neq 0$, because $E\binom{j_1 \ldots j_k}{\alpha_1 \ldots \alpha_k}$ and $E\binom{i_1 \ldots i_h}{\beta_1 \ldots \beta_h}$ are analytically regular and Eq. (4.14.23) is just a transcripton in other symbols of Eq. (4.14.6). However the general case, when some of the α's or β's may be zero, can be treated in the same way as that used to show the existence of the frequencies of visit, see Eqs. (4.14.24)-(4.14.26). mbe

It is useful to reinterpret Proposition 33 as follows.

Given $\boldsymbol{\xi} \in W$, consider the $(\mathcal{E}, \mathbf{t})$ history of $\boldsymbol{\xi}$, see Definition 18, §4.14, p.341. It is the sequence of $\mathbf{a} = (a_0, a_1, \ldots)$, $ai = 0, 1, \ldots, s$ such that

$$S_{it_1}(\boldsymbol{\xi}) \in E_{a_i}, \qquad i = 0, 1, \ldots. \tag{4.14.27}$$

The frequencies of Eq. (4.14.19) can be "computed" from the history $\mathbf{a}$ as

$$\mathbf{p}\begin{pmatrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{pmatrix}\!\Big|\boldsymbol{\xi}\Big) = \lim_{N\to\infty} \frac{1}{N} n_N \begin{pmatrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{pmatrix}\!\Big|\mathbf{a}\Big), \tag{4.14.28}$$

where $n_N(\ldots)$ is the number of values of h, integer and smaller than N, such that

$$S_{(h+j_1)t_1}(\boldsymbol{\xi}) \in E_{\alpha_a}, \ldots, S_{(h+j_k)t_1}(\boldsymbol{\xi}) \in E)a_k \tag{4.14.29}$$

[see Eqs. (4.14.15) and (4.14.19)], i.e., it is the number of times when

$$a_{h+j_1} = \alpha_1, \ldots, a_{h+j_k} = \alpha_k \tag{4.14.30}$$

occur simultaneously, with h integer in $[0, N)$. In other words, Eq. (4.14.28) says that $\mathbf{p}\binom{j_1 \ldots j_k}{\alpha_1 \ldots \alpha_k}\big|\boldsymbol{\xi}\big)$ is the frequency of appearance of the "string $\alpha_1, \ldots, \alpha_k$ at sites following each other at successive distances $j_2 - j_1, j_3 - j_2, \ldots, j_k - j_{k-1}$ in the history $\mathbf{a}$ of $\boldsymbol{\xi}$. It is then natural to set the following general definition.

19 Definition. *Let* $\mathbf{a} = (a_i)_{i=0}^{\infty}$ *be a sequence,* $a_i = 0, 1, \ldots, s, \forall\, i \in \mathcal{Z}_+$. *Given* $k > 0$, $0 \le j_1 < j_2 < \ldots < j_k$ *integers and* $\alpha_1, \ldots, \alpha_k$ *in* $(0, 1, \ldots, s)$

we say that a "string homologous to $\binom{j_1 \ldots j_k}{\alpha_1 \ldots \alpha_k}$*" is "realized in* $\mathbf{a}$ *at the* h*-th site" if*

$$a_{h+j_1} = \alpha_1, a_{h+j_2} = \alpha_2, \ldots, a_{h+j_k} = \alpha_k. \tag{4.14.31}$$

The frequency of realization in $\mathbf{a}$ *of strings homologous to* $\binom{j_1 \ldots j_k}{\alpha_1 \ldots \alpha_k}$ *will be defined in terms of the quantity*

$$\mathbf{p}_N \left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix} \middle| \mathbf{a} \right) = \frac{1}{N} \Big\{ \textit{number of times a string homologous} \\ \textit{to } \binom{j_1 \ldots j_k}{\alpha_1 \ldots \alpha_k} \textit{ is realized in } \mathbf{a} \textit{ at sites } h \textit{ between } 0 \textit{ and } N \Big\} \tag{4.14.32}$$

We shall set

$$\mathbf{p} \left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix} \middle| \mathbf{a} \right) \stackrel{def}{=} \lim_{N \to \infty} \mathbf{p}_N \left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix} \middle| \mathbf{a} \right) \tag{4.14.33}$$

whenever the limit exists.
We shall say that a sequence $\mathbf{a}$ *is "ergodic" if:*
(i) it has well-defined frequencies of appearance for all the strings of symbols, i.e., the limits (4.14.33) exist for all choices of the indices;
(ii) there are at least two distinct symbols α, β *occurring with positive frequency in* $\mathbf{a}$*:*

$$\mathbf{p}\left(\begin{matrix} 0 \\ \alpha \end{matrix} \middle| \mathbf{a} \right) > 0, \qquad \mathbf{p}\left(\begin{matrix} 0 \\ \beta \end{matrix} \middle| \mathbf{a} \right) > 0; \tag{4.14.34}$$

(iii) for all choices of indices

$$\lim_{N \to \infty} \frac{1}{N} \sum_{\ell=0}^{N-1} \mathbf{p} \left(\begin{matrix} j_1 \ldots j_k \, i_{1+\ell} \ldots i_{h+\ell} \\ \alpha_1 \ldots \alpha_k \, \beta_1 \quad \ldots \beta_h \end{matrix} \middle| \mathbf{a} \right) \\ = \mathbf{p} \left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix} \middle| \mathbf{a} \right) \mathbf{p} \left(\begin{matrix} i_1 \ldots i_h \\ \beta_1 \ldots \beta_h \end{matrix} \middle| \mathbf{a} \right). \tag{4.14.35}$$

As $k, j_1, \ldots, j_k, \alpha_1, \ldots, \alpha_k$ *vary, the family of numbers (4.14.33) will be called the "distribution of* $\mathbf{a}$*".*
If a only verifies (i) [or (i) and (ii)], it will be called a "sequence with well-defined frequencies" (respectively, a "sequence with nontrivial frequencies") of the occurrence of the symbols.
Finally, an ergodic sequence is said "mixing" if for all the choices of indices,

$$\lim_{\ell \to \infty} \mathbf{p} \left(\begin{matrix} j_1 \ldots j_k \, i_{1+\ell} \ldots i_{h+\ell} \\ \alpha_1 \ldots \alpha_k \, \beta_1 \quad \ldots \beta_h \end{matrix} \middle| \mathbf{a} \right) = \mathbf{p} \left(\begin{matrix} j_1 \ldots j_k \\ \alpha_1 \ldots \alpha_k \end{matrix} \middle| \mathbf{a} \right) \mathbf{p} \left(\begin{matrix} i_1 \ldots i_h \\ \beta_1 \ldots \beta_h \end{matrix} \middle| \mathbf{a} \right). \tag{4.14.36}$$

which is obviously stronger than Eq. (4.14.35).

Observations.
(1) From the definition, Eq. (4.14.33), of the distribution of $\mathbf{a}$ as a family of frequencies of certain events, it immediately follows that such numbers verify Eqs. (4.14.20), (4.14.21), and (4.14.22) with $\mathbf{a}$ replacing $\boldsymbol{\xi}$.
(2) Using the language of probability theory (see §2.23), we can say that to any sequence $\mathbf{a}$ with well-defined frequencies of occurrence of the symbols it is possible to associate a family $(\mathcal{E}_k, \mathbf{p}_k)_{k=1}^{\infty}$ of probability distributions as follows. $\mathcal{E}_k$ will be the set of $(s+1)^k$ events, which we can denote $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{k-1})$, $\alpha_i = 0, 1, \ldots, s$, whose probability is $\mathbf{p}\begin{pmatrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{pmatrix}$. By Eq. (4.14.33), this probability coincides, by definition, with the frequency of occurrence in $\mathbf{a}$ of strings homologous to $\begin{pmatrix} 0 \ldots & k-1 \\ \alpha_0 \ldots & \alpha_{k-1} \end{pmatrix}$.
For this reason, the sequence $(\mathcal{E}_k, \mathbf{p}_k)_{k=1}^{\infty}$ is also called the "probability distribution of the symbols of $\mathbf{a}$" and $\mathbf{p}\left(\begin{matrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{matrix} \middle| \mathbf{a}\right)$ is called the "probability of he string $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{k-1})$ in $\mathbf{a}$".

Proposition 33 can be reinterpreted in terms of the above definition:

34 Proposition. *By the assumptions of the preceding proposition, denote for* $\boldsymbol{\xi} \in W$ *the* $(\mathcal{E}, \mathbf{t})$ *history,* $\mathbf{t} = (i\, t_1)_{i=0}^{\infty}$, $t_1 > 0$ *of* $\boldsymbol{\xi}$ *as* $\mathbf{a}(\boldsymbol{\xi})$*. Then, if Eq. (4.14.34) holds,* $\mathbf{a}(\boldsymbol{\xi})$ *is an ergodic non mixing sequence.*

Observation. The only statement not already contained in Proposition 33 is the one concerning mixing.

Proof. By the assumed analytic integrability of the system, we can imagine that $\mathbf{a} = \mathbf{a}(\boldsymbol{\xi})$ is the $(\mathcal{E}, \mathbf{t})$ history of a point $\boldsymbol{\varphi} \in \mathcal{T}^\ell$ with respect to an analytically regular partition $\mathcal{E} = (E_0, \ldots, E_p)$ of $\mathcal{T}^\ell$ and to the transformations $(S_t)_{t \in \mathcal{R}}$ of $\mathcal{T}^\ell$ given by

$$S_t \boldsymbol{\varphi} = \boldsymbol{\varphi} + t\, \boldsymbol{\omega} \mod 2\pi. \tag{4.14.37}$$

For simplicity, we shall only deal with the case when $\omega_1, \ldots, \omega_\ell, \sigma = \frac{2\pi}{t_1}$ are rationally independent and when it is also assumed that there are two sets E_α, E_β such that $\mathbf{p}\left(\begin{matrix} 0 \\ \alpha \end{matrix} \middle| \mathbf{a}\right) > 0$, $p\left(\begin{matrix} 0 \\ \beta \end{matrix} \middle| \mathbf{a}\right) > 0$, having a diameter so small that there is a point $\boldsymbol{\varphi}_0 \in E_\alpha$ at a distance from E_β, greater than twice the diameter of E_α.

These are serious restrictions. However, the general case can be reduced to the above, as it will become apparent after having gone through the problems at the end of this section.

The rational independence assumption of $\omega_1, \ldots, \omega_\ell, \sigma$ and the analytic regularity of $\mathcal{E}$ imply that

$$\mathbf{p}\left(\begin{matrix} 0 & j \\ \gamma & \gamma' \end{matrix} \middle| \mathbf{a}\right) = \frac{1}{(2\pi)^\ell} \int_{E_\xi \cap S_{-j\, t_1} E_\beta} d\boldsymbol{\varphi}, \qquad \forall\, \gamma, \gamma' \tag{4.14.38}$$

[see Proposition 30, p.342, Eq. (4.14.5)].

If $\mathbf{a}(\boldsymbol{\xi})$ were mixing, by Eq. (4.14.36), one would also have

$$\lim_{j\to\infty} \mathbf{p}\begin{pmatrix} 0 & j \\ \gamma & \gamma' \end{pmatrix}\!\Big|\mathbf{a}\Big) = \mathbf{p}\begin{pmatrix} 0 \\ \alpha \end{pmatrix}\!\Big|\mathbf{a}\Big)\, V p\begin{pmatrix} 0 \\ \beta \end{pmatrix}\!\Big|\mathbf{a}\Big) > 0 \tag{4.14.39}$$

However this would mean that for j large enough, it should be

$$\mathbf{p}\begin{pmatrix} 0 & j \\ \alpha & \beta \end{pmatrix}\!\Big|\mathbf{a}\Big) = \int_{E_\xi \cap S_{-j\,t_1} E_\beta} d\boldsymbol{\varphi} > 0 \tag{4.14.40}$$

Hence, $E_\alpha \cap S_{-j\,t_1} E_\beta \neq \emptyset$ eventually. But, by the rational independence of $\omega_1, \ldots, \omega_\ell, \sigma$ the trajectory $\widetilde{\boldsymbol{\varphi}} - j\,t_1\,\boldsymbol{\omega}$, $j \geq j_0$, of any point $\widetilde{\boldsymbol{\varphi}}$ chosen in E_β is dense in $\mathcal{T}^\ell$ (see §4.2). Therefore, given $\boldsymbol{\varphi}_0 \in E_\alpha$, there must exist infinitely many values of $j \geq 0$ such that the distance of $\widetilde{\boldsymbol{\varphi}} - j\,t_1\,\boldsymbol{\omega}$ from $\boldsymbol{\varphi}_0$ is less than the diameter of E_β. For such values of j, it must be that $E_\alpha \cap S_{-j\,t_1} E_\beta = \emptyset$ since these torus rotations do not deform the sets but they only translate them, and $\boldsymbol{\varphi}_0$ is chosen so that $d(\boldsymbol{\varphi}_0, E_\beta) > \{$twice the diameter of $E_\beta\}$. mbe

4.14.1 Exercises and Problems

Solve the following connected sequence of problems for $\ell = 2$ first drawing graphical representations of the various maps and transformations. The notations are those of §4.14. The aim is to solve problem 8 below.

1. Let $\omega_1, \ldots, \omega_\ell$ be rationally dependent and not all zero. Show that there exists $\overline{\ell} < \ell$ rationally independent numbers $\widehat{\omega}_1, \ldots, \widehat{\omega}_{\overline{\ell}}$ and an $\ell \times \overline{\ell}$ matrix J with integer coefficients and such that $\boldsymbol{\omega} = J\widetilde{\boldsymbol{\omega}}$, i.e., $\omega_j = \sum_{k=1}^{\overline{\ell}} J_{jk}\widehat{\omega}_k$, $j = 1, \ldots, \ell$.

2. In $\mathcal{R}^\ell$ consider the plane $\pi_\ell = J\mathcal{R}^\ell = \{\mathbf{x}\,|\,x_j = \sum_{k=1}^{\overline{\ell}} J_{jk}y_k,\ \mathbf{y} \in \mathcal{R}^{\overline{\ell}}$ and the plane $\pi_\ell^\perp$ orthogonal to it. Show that there exists an $\ell \times (\ell - \overline{\ell})$ matrix $J^\perp$ with integer coefficients such that $\pi_\ell^\perp = J^\perp \mathcal{R}^{\ell - \overline{\ell}}$.

3. Define the map $(J \times J^\perp)_T$ of $\mathcal{T}^{\overline{\ell}} \times \mathcal{T}^{\ell-\overline{\ell}}$ onto $\mathcal{T}^\ell$, $\forall\ (\boldsymbol{\vartheta}, \boldsymbol{\nu}) \in \mathcal{T}^{\overline{\ell}} \times \mathcal{T}^{\ell-\overline{\ell}}$ as:

$$(J \times J^\perp)(\boldsymbol{\vartheta}, \boldsymbol{\nu}) = (J\boldsymbol{\vartheta} + J^\perp\boldsymbol{\nu} \mod 2\pi.$$

If one defines $\widehat{E} = (J \times J^\perp)_T^{-1} E$, for $E \subset \mathcal{T}^\ell$, and if E is analytically regular in $\mathcal{T}^\ell$, show that $\widehat{E}$ is such in $\mathcal{T}^{\overline{\ell}} \times \mathcal{T}^{\ell-\overline{\ell}} \equiv \mathcal{T}^\ell$. (*Hint:* Note that $(J \times J^\perp)_T \mathcal{T}^\ell$ regarded as a matrix denoted $(J \times J^\perp)_T$ linearly maps $\mathcal{R}^\ell$ onto $\mathcal{R}^\ell$; hence, $\det(J \times J^\perp) \neq 0$. Hence, $(J \times J^\perp)_T^{-1})E$ is analytically regular in $\mathcal{R}^\ell$ and $\widehat{E}$ is obtained by considering $(J \times J^\perp)_T^{-1}E$, after reducing mod 2π, the coordinates of its points, as a subset of the torus $\mathcal{T}^{\overline{\ell}} \times \mathcal{T}^{\ell-\overline{\ell}} \equiv \mathcal{T}^\ell$.)

4. If $\boldsymbol{\varphi}_0 \in \mathcal{T}^\ell$ and $\boldsymbol{\varphi}_0 = (J \times J^\perp)_T(\boldsymbol{\vartheta}_0, \boldsymbol{\nu}_0)$ show that the frequency of visits to E of the trajectory of $\boldsymbol{\varphi}_0$ under the transformation $\boldsymbol{\varphi}_0 \to \boldsymbol{\varphi}_0 + t\boldsymbol{\omega}$ coincides with the frequency of visit to $\widehat{E}$ of the trajectory of $(\boldsymbol{\vartheta}_0, \boldsymbol{\nu}_0)$ under the transformation $(\boldsymbol{\vartheta}_0, \boldsymbol{\nu}_0) \to (\boldsymbol{\vartheta}_0 + \widetilde{\boldsymbol{\omega}}\,t, \boldsymbol{\nu}_0)$ (*Hint:* Note that $\boldsymbol{\varphi}_0 + \boldsymbol{\omega}\,t = (J \times J^\perp)_T(\boldsymbol{\vartheta}_0 + \widehat{\boldsymbol{\omega}}\,t), \boldsymbol{\nu}_0)$ by the construction of J.)

5. Let $\widehat{E}(\boldsymbol{\nu}_0) = \widehat{E} \cap \{(\boldsymbol{\vartheta}, \boldsymbol{\nu})\,|\,(\boldsymbol{\vartheta}, \boldsymbol{\nu}) \in \mathcal{T}^{\overline{\ell}} \times \mathcal{T}^{\ell-\overline{\ell}},\ \boldsymbol{\nu} = \boldsymbol{\nu}_0\}$, then the frequency of visits to E of the trajectory of $\boldsymbol{\varphi}_0$ for the transformations $\boldsymbol{\varphi}_0 \to \boldsymbol{\varphi}_0 + t\,\boldsymbol{\omega}$ coincides with the frequency of visits to $\widehat{\widehat{E}}(\boldsymbol{\nu}_0) = \{\,\boldsymbol{\vartheta}\,|\,\boldsymbol{\vartheta} \in \mathcal{T}^{\widetilde{\ell}},\ (\boldsymbol{\vartheta}, \boldsymbol{\nu}_0) \in \widehat{E}(\boldsymbol{\nu}_0)\} \subset \mathcal{T}^{\widetilde{\ell}}$ by the trajectory of $\boldsymbol{\vartheta}_0$ under the

transformation $\boldsymbol{\vartheta}_0 \to \boldsymbol{\vartheta}_0 + \widehat{\boldsymbol{\omega}}\, t$. Furthermore, if E is analytically regular in $\mathcal{T}^\ell$, then $\widehat{\widehat{E}}(\boldsymbol{\nu}_0)$ is such in $\mathcal{T}^\ell$ (*Hint:* Interpret $\widehat{\widehat{E}}(\boldsymbol{\nu}_0)$ as the intersection of $\widehat{E}(\boldsymbol{\nu}_0)$ with a "plane".)

6. If $\omega_1, \ldots, \omega_\ell$ are rationally independent but $\omega_1, \ldots, \omega_\ell, \sigma = \frac{2\pi}{t_1}$ are not rationally independent there are ℓ integers $m_1, \ldots, m_\ell$ and $q > 0$, integer too, such that

$$\sigma = \frac{\mathbf{m} \cdot \boldsymbol{\omega}}{q}.$$

The problem of the determination of the frequency of visits to $E \subset \mathcal{T}^\ell$ by the trajectory of $\boldsymbol{\varphi} \in \mathcal{T}^\ell$ under the map $\boldsymbol{\varphi} \to \boldsymbol{\varphi} + \boldsymbol{\omega}\frac{2\pi\, j}{\sigma}$, $j = 0, 1, \ldots$ is equivalent (via a suitable change of coordinates) to the analogous problem when the relation between σ and $\boldsymbol{\omega}$ is simply $\sigma = \frac{m}{q}\omega_1$. (*Hint:* The transformation is analogous to that described in Problems 2 and 3 above. It is the transformation associated, in the same way as above, to the matrix J of the transformation

$$\omega_1 = \omega_1' - \sum_{i=2}^{\ell} m_i \omega_i', \quad \omega_j = m_1 \omega_j', \qquad j = 2, \ldots, \ell.)$$

7. Consider the trajectory of $\boldsymbol{\varphi}_0 \in \mathcal{T}^\ell$ under the transformations $\boldsymbol{\varphi}_0 \to \boldsymbol{\varphi}_0 + \frac{2\pi}{\sigma}\, j\, \boldsymbol{\omega}$ with $\sigma = \frac{m}{q}\omega_1$, m, q integers, and assume that $\omega_1, \ldots, \omega_\ell$ are rationally independent.
Think of $\mathcal{T}^\ell$ as $\mathcal{T}^1 \times \mathcal{T}^{\ell-1}$ and, if $(\varphi, \boldsymbol{\psi}) \in \mathcal{T}^1 \times \mathcal{T}^{\ell-1}$, show that the map under analysis can be written as $(\varphi, \boldsymbol{\psi}) \to (\varphi + \frac{2\pi q}{m} j, \boldsymbol{\psi} + \boldsymbol{\omega}' j)$ where $\boldsymbol{\omega}' = (\omega_2', \ldots, \omega_\ell')$ are $\ell - 1$ rationally independent numbers which together with 2π form a set of ℓ rationally independent numbers.
If $E \subset \mathcal{T}^\ell$ is analytically regular, show that the frequency of visit to E exists and depends only on φ. (*Hint:* Note that

$$\begin{aligned}
&\frac{1}{Mm} \sum_{j=0}^{Mm-1} \chi_E(\varphi_0 + \frac{2\pi q}{m} j, \boldsymbol{\psi}_0 + \boldsymbol{\omega}'\, j) \\
&= \frac{1}{Mm} \sum_{k=0}^{m-1} \sum_{p=0}^{M-1} \chi_E(\varphi_0 + \frac{2\pi q}{m}(k + mp), (\boldsymbol{\psi}_0 + k\boldsymbol{\omega}') + mp\boldsymbol{\omega}') \\
&= \frac{1}{m} \sum_{k=0}^{m-1} \Big(\frac{1}{M} \sum_{p=0}^{M-1} \chi_E(\varphi_0 + \frac{2\pi q}{m}(k + mp), (\boldsymbol{\psi}_0 + k\boldsymbol{\omega}') + mp\boldsymbol{\omega}') \Big),
\end{aligned}$$

and, letting $\varphi_k = \varphi_0 + \frac{2\pi q}{m} k$, $\boldsymbol{\psi}_k = \boldsymbol{\psi}_0 + k\, \boldsymbol{\omega}'$, this can be rewritten

$$\frac{1}{m} \sum_{k=0}^{m-1} \Big(\frac{1}{M} \sum_{p=0}^{M-1} \chi_E(\varphi_k, \psi_k + mp\boldsymbol{\omega}') \Big) = \frac{1}{m} \sum_{k=0}^{m-1} \Big(\frac{1}{M} \sum_{p=0}^{M-1} \chi_{E_k(\varphi_0)}(\psi_k + mp\boldsymbol{\omega}') \Big),$$

where $E_k(\varphi_0) = \{\boldsymbol{\psi} \,|\, \boldsymbol{\psi} \in \mathcal{T}^{\ell-1}, (\varphi_0 + \frac{2\pi q}{m} k, \boldsymbol{\psi}) \in E\}$ is still analytically regular for $k = 0, \ldots, m-1$. Hence, the frequency of visit to E exists because $m\boldsymbol{\omega}'$ has rationally independent components and it is given by $\frac{1}{m} \sum_{k=0}^{m-1} \int_{E_k(\varphi_0)} \frac{d\boldsymbol{\varphi}'}{(2\pi)^{\ell-1}}$.)

8. On the basis of the above problems, deduce the proofs of Propositions 30 and 31 in the general case, from their validity in the rationally independent cases.

4.15 Analytic Integrability Criteria. Complexity of Motions and Entropy

Summarizing the preceding sections discussion, the following criteria of non analytic integrability, on a phase space subset W, have been obtained for an analytic time-independent Hamiltonian system:

(i) if in W there is one $\boldsymbol{\xi}$ whose $(\mathcal{E},\mathbf{t})$ history ($\mathbf{t}=(i\,t_1)_{i=0}^{\infty}$) on an analytically regular partition $\mathcal{E}$ of W contains some strings without well-defined frequency of occurrence;
(ii) if in W there is one $\boldsymbol{\varphi}$ whose trajectory T has a closure $\overline{T}$ that cannot be mapped bicontinuously on a torus $\mathcal{T}^s$, $s\le\ell$, ℓ being the number of degrees of freedom;
(iii) if in W there is one $\boldsymbol{\xi}$ whose $(\mathcal{E},\mathbf{t})$ history on an analytically regular partition $\mathcal{E}$ of W has nontrivial frequency distributions but is not ergodic;
(iv) if in W there is one $\boldsymbol{\xi}$ whose $(\mathcal{E},\mathbf{t})$ history on an analytically regular partition $\mathcal{E}$ of W is "too ergodic", i.e., mixing.

The review of non integrability criteria will be concluded by examining another very interesting property of the analytically integrable systems: namely that the motions of such systems have a "small complexity". This leads to another non integrability criterion, see (v), p.359.

To obtain such a result a quantitative meaning is needed for the notion of "complexity" of the motions associated with points moving on a regular (analytic) surface under the action of a family ("semigroup") $(S_t)_{t\in\mathcal{R}_+}$ of C^∞ (analytic) transformations.

A natural way to evaluate the complexity of a motion is to count the number of different strings of history appearing in the $(\mathcal{E},\mathbf{t})$ history of the motion on an analytically regular partition.

20 Definition. *Let* $\mathbf{a}$ *be a sequence* $\mathbf{a}=(a_i)_{i=0}^{\infty}$, $a_i\in(0,\ldots,s)$. *Assume that* $\mathbf{a}$ *has well-defined frequencies of symbol appearances (see p.348).*
The "number of strings of symbols of length k *appearing in* $\mathbf{a}$*" is defined as*

$$\mathcal{N}_{abs}(\mathbf{a},k)=\Big\{\textit{number of choices of } (\alpha_0,\ldots,\alpha_{k-1})\in(0,\ldots,s)^k \textit{ such that } \mathbf{p}\left(\begin{matrix}0&\ldots&k-1\\ \alpha_0&\ldots&\alpha_{k-1}\end{matrix}\,\Big|\,\mathbf{a}\right)>0\Big\} \tag{4.15.1}$$

where, we recall, $\mathbf{p}\left(\begin{smallmatrix}0&\ldots&k-1\\ \alpha_0&\ldots&\alpha_{k-1}\end{smallmatrix}\,\big|\,\mathbf{a}\right)$ *denotes the frequency of appearance in* $\mathbf{a}$ *of a string homologous to* $\left(\begin{smallmatrix}0&\ldots&k-1\\ \alpha_0&\ldots&\alpha_{k-1}\end{smallmatrix}\,\big|\,\mathbf{a}\right)$*, see Definition 19, p.348*

Clearly $\mathcal{N}_{abs}s(\mathbf{a},k)<(s+1)^k$. *We shall set*[18]

$$S_{abs}(\mathbf{a})=\lim_{k\to+\infty}\frac{1}{k}\log\mathcal{N}_{abs}(\mathbf{a},k) \tag{4.15.2}$$

which we call the "absolute complexity" of the sequence $\mathbf{a}$.

[18] The limit always exists (see Problem 21, p.364).

Observations.
(1) The number in Eq. (4.15.2) can give an idea of how complex the sequence a might be. However, $S_{abs}(\mathbf{a})$ is a rather rough measure of the complexity of $\mathbf{a}$: in its evaluation, in fact one puts on the same footing strings occurring in $\mathbf{a}$ with a frequency of occurrence much smaller than that of other strings or, by the Observation (2), to Definition 19, p.348, with a "probability" much smaller than that of others.
(2) The existence of the limit of Eq. (4.15.8) is easy to prove and very instructive (see Problem 23 below).

The following more sophisticated definition takes into account the possibility that some strings may be present in a with extremely small probability and gives them less importance.

21 Definition. *Let* $\mathbf{a} = (a_i)_{i\in\mathcal{Z}_+}$, $\mathbf{a} = (a_i)_{i=0}^{\infty}$, $a_i = 0, 1, \ldots, s$ *be a sequence with well defined frequencies of symbol occurrence as in Definition 20.*
Given $\varepsilon > 0$, *consider all the possible subsets* $\mathcal{C}_\varepsilon$ *of the set of the* k*-tuples* $\alpha_0, \ldots, \alpha_{k-1}$, $\alpha_i = 0, \ldots, s$, *such that*

$$\sum_{\alpha_0,\ldots,\alpha_{k-1}} \mathbf{p}\left(\begin{matrix} 0 & \ldots & k-1 \\ \alpha_0 & \ldots & \alpha_{k-1} \end{matrix} \middle| \, \mathbf{a} \right) < \varepsilon. \tag{4.15.3}$$

These are the sets $\mathcal{C}_\varepsilon$ *of "*k*-strings" (strings of length* k*) whose total frequency of occurrence is smaller than* ε*. Let*

$$\mathcal{N}(\mathbf{a}, k, \varepsilon) \stackrel{def}{=} \text{minimum, over the choices of } \mathcal{C}_\varepsilon, \text{ of the number of } k\text{-tuples outside } \mathcal{C}_\varepsilon \tag{4.15.4}$$

and let

$$S(\mathbf{a}, \varepsilon) = \limsup_{k\to+\infty} \frac{1}{k} \log \mathcal{N}(\mathbf{a}, k, \varepsilon), \tag{4.15.5}$$

$$S(\mathbf{a}) = \lim_{\varepsilon\to 0} S(\mathbf{a}, \varepsilon). \tag{4.15.6}$$

This last quantity will be called the "entropy" of $\mathbf{a}$ *and it can also be regarded as a measure of the complexity of* $\mathbf{a}$*.*

Observations.
(1) This is a measure of complexity more interesting than Eq. (4.15.2). Through Eq. (4.15.4) and the two limits in Eqs. (4.15.5) and (4.15.6), in some way, one discards from the number of strings of $\mathbf{a}$ those which appear with a very small frequency (see, also, Proposition 37 to follow).
(2) Obviously,

$$0 < S(\mathbf{a}) \le S_{abs}(\mathbf{a}) \le \log(s+1), \tag{4.15.7}$$

and one can note that the two numbers given in Eqs. (4.15.2) and (4.15.6) can be thought of as obtained by permuting the following two limits:

$$S_{abs}(\mathbf{a}) = \lim_{k\to\infty} \lim_{\varepsilon\to 0} \frac{1}{k} \log \mathcal{N}(\mathbf{a}, k, \varepsilon), \tag{4.15.8}$$

$$S(\mathbf{a}) = \lim_{\varepsilon\to 0} \lim_{k\to\infty} \frac{1}{k} \log \mathcal{N}(\mathbf{a}, k, \varepsilon), \tag{4.15.9}$$

if all the above limits exist.
(3) The term entropy given to Eq. (4.15.6) is due to the analogy of this definition with Boltzmann's fundamental idea on the proportionality between the entropy of the state of a system, in the thermodynamic sense of the word, and the number of ways of realizing the same macroscopic state by equivalent microscopic states.
This analogy is evident if one is not biased by the various limit steps taken in Eqs. (4.15.5), (4.15.6), (4.15.8), and (4.15.9), and at first one ignores them.

The following proposition holds.

35 Proposition. *Consider a Hamiltonian system analytically integrable on the phase-space subset W. Let $\mathcal{E} = (E_0, \ldots, E_s)$ be an analytically regular partition of W. Let $t_1 > 0$, $\mathbf{t} = (i\, t_1)_{i=0}^{\infty}$.*
For all $\boldsymbol{\xi} \in W$, denote $\mathbf{a}(\boldsymbol{\xi})$ the $(\mathcal{E}, \mathbf{t})$ history of $\boldsymbol{\xi}$. Then

$$S(\mathbf{a}(\boldsymbol{\xi})) = 0, \qquad \forall\, \boldsymbol{\xi} \in W. \tag{4.15.10}$$

Observation. As already seen in the propositions of §4.14, the statement of this proposition is an immediate consequence of an analogous proposition concerning the torus rotations. In this case, the proposition is the following.

36 Proposition. *Let $\boldsymbol{\omega} \in \mathcal{R}^\ell$ and let $(S_t)_{t\in\mathcal{R}}$ be the quasi-periodic flow on $\mathcal{T}^\ell$ with pulsations $\boldsymbol{\omega}$ (i.e. $S_t\boldsymbol{\varphi} = \boldsymbol{\varphi} + t\boldsymbol{\omega}$). Consider the transformations $(S_{jt_1})_{j=0}^{\infty}$, $t_1 > 0$, and let $\mathcal{E} = (E_0, \ldots, E_s)$ be an analytically regular partition of $\mathcal{T}^\ell$ into $(s+1)$ sets. The $(\mathcal{E}, \mathbf{t})$-history $\mathbf{a}(\boldsymbol{\varphi})$ of $\boldsymbol{\varphi} \in \mathcal{T}^\ell$ is such that*

$$S(\mathbf{a}(\boldsymbol{\varphi})) = 0, \qquad \forall\, \boldsymbol{\varphi} \in \mathcal{T}^\ell. \tag{4.15.11}$$

Observations.
(1) The argument presented in the proof below essentially gives the proof of a more general theorem of great importance in the theory of entropy ("Koushnirenko's theorem", [4]).
(2) Actually, one could prove a stronger result namely,

$$S_{abs}(\mathbf{a}(\boldsymbol{\varphi})) = 0, \qquad \forall\, \boldsymbol{\varphi} \in \mathcal{T}^\ell. \tag{4.15.12}$$

However, in the course of the proof, we show Eq. (4.15.12) only in the $\ell = 1$ case. The argument could be adapted to prove Eq. (4.15.12) in general. However, for $\ell > 1$, an alternative proof of the weaker result of Eq. (4.15.11) is preferable because the method of this proof is in itself interesting and, as mentioned in Observation (1), contains the germs of interesting extensions.
(3) Equations (4.15.10) and (4.15.11) have an interesting monotonicity property: if $\mathcal{E}'$ is a partition finer than $\mathcal{E}$ in the sense that every set in $\mathcal{E}$ can be

thought of as a union of sets in $\mathcal{E}$, then the absolute complexity (and the entropy) of $\mathbf{a}'(\boldsymbol{\varphi})$ is not smaller than that of $\mathbf{a}(\boldsymbol{\varphi})$. This reflects the intuitively clear fact that by increasing the precision of the measurements, the motion can only look more complicated since more of its features may become manifest.

PROOF: As mentioned in Observation (2), the cases $\ell = 1$ and $\ell > 1$ will be considered separately . We only treat the case when $\omega_1, \ldots, \omega_\ell, \frac{2\pi}{t_1}$ are rationally independent. The problems of §4.14 show that the general case can be reduced to this special one.

Case $\ell = 1$. To fix the ideas, suppose $s = 1$ and $E_0 = (\lambda, 2\pi)$, $E_1 = [0, \lambda]$, $\lambda \in (0.2\pi)$. Consider the images of the points 0 and λ for the maps $\boldsymbol{\varphi} \to \boldsymbol{\varphi} + jt_1\omega_1$, $j = 0, \ldots, k-1$ There are at most $2(k+1)$ points (and at least 2) dividing the interval $[0, 2\pi]$ in $2(k+1)$, at most, consecutive intervals $J_1, J_2, \ldots$. Then all points internal to some such interval have the same $(\mathcal{E}, \mathbf{t})$ history in the first k sites of their history.

To the $2(k+1)$, at most, histories of the points internal to the above intervals, we can add the $2(k+1)$ histories, at most, of their extreme points. We thus obtain all the possible strings of the history with length k that can appear in the $(\mathcal{E}, \mathbf{t})$ history of a point $\varphi \in \mathcal{T}^1$. Hence,

$$\mathcal{N}_{abs}(\mathbf{a}(\varphi), k) \le 4(k+1) \tag{4.15.13}$$

and Eq. (4.15.12) follows from the definition given by Eq. (4.15.2).

Case $\ell > 1$. The entire proof will be based on the possibility of estimating the volume $|E|$ of a set E in terms of the area $|\partial E|$ of its boundary ∂E. If $E \subset \mathcal{R}^\ell$ is a bounded set its volume $|E|$ cannot exceed the volume of the sphere with surface area equal to the surface area $|\partial E|$ of E ("isoperimetric inequality"). So an inequality of the type

$$|E| < C_\ell \, |\partial E|^{\frac{\ell}{\ell-1}} \tag{4.15.14}$$

holds C_ℓ being a suitable E-independent constant. However, on $\mathcal{T}^\ell$, such an inequality is false for sets which "wrap around $\mathcal{T}^\ell$" (e.g., if $E = \mathcal{T}^\ell$, $|E| = (2\pi)^\ell$, $|\partial E| = 0$ as $\partial E = \emptyset$); but of course, it is still true for sets with small enough diameter.

To apply isoperimetric inequalities in $\mathcal{T}^\ell$, it is therefore useful to think of $\mathcal{T}^\ell$ as the union of many small sets. We shall regard $\mathcal{T}^\ell$ as a union of 2^ℓ cubes with side π parameterized by an index σ:

$$C_\sigma = \{\boldsymbol{\varphi} \,|\, \boldsymbol{\varphi} \in \mathcal{T}^\ell,\, \pi\sigma_i \le \varphi_i \le \pi\,(\sigma_i + 1),\, i = 1, \ldots, \ell\} \tag{4.15.15}$$

where each σ_i takes the value 0 or 1. We call Σ the set of the 2^ℓ σ's.

Given $(\alpha_0, \ldots, \alpha_k) \in \{0, \ldots, s\}^k$ and $(\sigma_0, \ldots, \sigma_{k-1}) \in \Sigma^k$, consider the sets

$$E\begin{pmatrix}0\,\ldots\,k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix} = E_{\alpha_0}\cap S_{-t_1}E_{\alpha_1}\cap\ldots\cap S_{-(k-1)t_1}E_{\alpha_{k-1}}$$
$$B\begin{pmatrix}0\,\ldots\,k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix} = C_{\sigma_0}\cap S_{-t_1}C_{\sigma_1}\cap\ldots\cap S_{-(k-1)t_1}C_{\sigma_{k-1}} \tag{4.15.16}$$

Since the rotations of the torus are "rigid transformations", i.e., they do not change the form and volume of the sets that they transform, it will be possible to infer that the sum of the surfaces of the sets $E\cap B$, with E, B like Eq. (4.15.16) with the same value of k, is such that

$$\sum_{\substack{\alpha_0,\ldots,\alpha_{k-1}\\ \sigma_0,\ldots,\sigma_{k-1}}}\left|\partial\left(E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\cap B\begin{pmatrix}0\ldots k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix}\right)\right|\le 2(k+1)L, \tag{4.15.17}$$

where $L=\sum_{j=0}^{s}|\partial E_j|+2^\ell(2\ell\pi^{\ell-1})$. This simple relation follows from the geometric observation that

$$\bigcup_{\substack{\alpha_0,\ldots,\alpha_{k-1}\\ \sigma_0,\ldots,\sigma_{k-1}}}\partial\Big(E\begin{pmatrix}0\,\ldots\,k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\cap B\begin{pmatrix}0\,\ldots\,k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix}\Big)$$
$$=\bigcup_{h=0}^{k-1}\big[S_{-ht_1}(\partial E_{\alpha_h})\cup S_{-ht_1}(\partial C_{\sigma_h})\big], \tag{4.15.18}$$

and the right-hand-side points are counted twice in the left-hand side except for a subset of total area zero corresponding to the edges and corners of the sets $E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\cap B\begin{pmatrix}0\ldots k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix}$. We can now use Eq. (4.15.14) to bound

$$\mathbf{p}\begin{pmatrix}0\,\ldots\,k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\Big)=\int_{E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}}\frac{d\boldsymbol{\varphi}}{(2\pi)^\ell}=\frac{E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}}{(2\pi)^\ell}$$
$$=\sum_{\sigma_0,\ldots,\sigma_\ell}\frac{\left|E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\cap B\begin{pmatrix}0\ldots k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix}\right|}{(2\pi)^\ell}$$
$$\le C_\ell\sum_{\sigma_0,\ldots,\sigma_\ell}\frac{\left|\partial\Big(E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\cap B\begin{pmatrix}0\ldots k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix}\Big)\right|^{\frac{\ell}{\ell-1}}}{(2\pi)^\ell} \tag{4.15.19}$$
$$\le C_\ell\left(\sum_{\sigma_0,\ldots,\sigma_\ell}\frac{\left|\partial\Big(E\begin{pmatrix}0\ldots k-1\\ \alpha_0\ldots\alpha_{k-1}\end{pmatrix}\cap B\begin{pmatrix}0\ldots k-1\\ \sigma_0\ldots\sigma_{k-1}\end{pmatrix}\Big)\right|}{(2\pi)^\ell}\right)^{\frac{\ell}{\ell-1}},$$

having used the rational independence of $(\omega_1,\ldots,\omega_\ell,\frac{2\pi}{t_1})$ in the first step [applying Proposition 30, Eq. (4.14.5)], and in the last step the inequality

$(\alpha+\beta)^x \ge \alpha^x+\beta^x$, $\forall\, x \ge 1, \forall\, \alpha,\beta \ge 0$, has also been used. The isoperimetric inequality has been used in the intermediate step.

Equations (4.15.19) and (4.15.17) will now be used to estimate the total frequency of the strings of length k in $\mathbf{a}(\boldsymbol{\varphi})$ having "small probability" and, "precisely" such that given $\eta > 0$,

$$\mathbf{p}\begin{pmatrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{pmatrix}\!\Big|\mathbf{a}(\boldsymbol{\varphi})\Big) \le e^{-k\eta}. \tag{4.15.20}$$

Recalling the ideas involved in the proof of the Chebysčev inequality, Proposition 34, p.119, and if the label $*$ indicates that the sum is restricted to $\mathbf{p}\left(\begin{smallmatrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{smallmatrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\right) \le e^{-k\eta}$,

$$\begin{aligned} &\sum^{*}_{\alpha_0,\ldots,\alpha_{k-1}} \mathbf{p}\left(\begin{matrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{matrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\right) \\ &\le \sum_{\alpha_0,\ldots,\alpha_{k-1}} \Big(\frac{e^{-\eta\, k}}{\mathbf{p}\left(\begin{smallmatrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{smallmatrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\right)}\Big)^{\gamma} \mathbf{p}\left(\begin{matrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{matrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\right) \\ &\le \sum_{\alpha_0,\ldots,\alpha_{k-1}} \mathbf{p}\left(\begin{matrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{matrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\right)^{1-\gamma} \end{aligned} \tag{4.15.21}$$

no matter how $\gamma > 0$ is chosen. Then let $\gamma = \frac{1}{\ell}$, i.e., such that $(1-\gamma)\frac{\ell}{\ell-1} = 1$ and deduce from Eqs. (4.15.21), (4.15.19), and (4.15.17) that the total probability that Eq. (4.15.20) holds is bounded by

$$\begin{aligned} &e^{-\frac{\eta\, k}{\ell}}\Big(\frac{C_\ell}{(2\pi)^\ell}\Big)^{1-\gamma} \sum_{\substack{\alpha_0,\ldots,\alpha_{k-1} \\ \sigma_0,\ldots,\sigma_{k-1}}} \Big|\partial\Big(E\left(\begin{smallmatrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{smallmatrix}\right) \cap B\left(\begin{smallmatrix} 0 \ldots k-1 \\ \sigma_0 \ldots \sigma_{k-1} \end{smallmatrix}\right)\Big)\Big| \\ &\le (2\pi)^{1-\gamma} e^{-\frac{\eta\, k}{\ell}} 2(k+1)L \end{aligned} \tag{4.15.22}$$

Hence, given $\varepsilon > 0$ and $\eta > 0$, Eq. (4.15.22) shows that if k is so large that the right-hand side of Eq. (4.15.20) is smaller than ε, we can find, among the sets $\mathcal{C}_\varepsilon$ appearing in Eq. (4.15.3), a set

$$\mathcal{C}_\varepsilon(\eta) = \{\text{set of the } k\text{-tuples } \alpha_1, \ldots \alpha_k, \text{ verifying Eq. (4.15.20)}\}. \tag{4.15.23}$$

Since [see Eq. (4.15.22) and Observation (1) to Definition 19, p.348] it is $\sum_{\alpha_0,\ldots,\alpha_{k-1}} \mathbf{p}\left(\begin{smallmatrix} 0 \ldots k-1 \\ \alpha_0 \ldots \alpha_{k-1} \end{smallmatrix}\Big|\mathbf{a}(\boldsymbol{\varphi})\right) = 1$ it becomes clear that the complement of $\mathcal{C}_\varepsilon(\eta)$ cannot contain more than $e^{\eta k}$ elements because it consists of sets with probability $\ge e^{-\eta k}$. One then finds that

$$\mathcal{N}(\varepsilon, \mathbf{a}(\boldsymbol{\varphi}), k) \le e^{\eta k} \tag{4.15.24}$$

if k is large enough. Hence,

$$S(\mathbf{a}(\boldsymbol{\varphi}),\varepsilon) \le \eta \tag{4.15.25}$$

and Eq. (4.15.11) follows from the arbitrariness of η.

So far the analytic regularity of $\mathcal{E}$ has been only used to deduce the first of Eqs. (4.15.19) which, as remarked elsewhere (see Observation (1), to Proposition 30, p.342), follows simply from the Riemann measurability of $E_1,\ldots,E_s$. However, in the general case when $\omega_1,\ldots,\omega_\ell,\frac{2\pi}{t_1}$ are not rationally independent, as assumed above, analytic regularity has to be used again to reduce the general case to the above-treated rationally independent case. mbe

The above propositions provide a further non integrability criterion.

(v) If in W there is one $\boldsymbol{\xi}$ whose $(\mathcal{E},\mathbf{t})$ history on an analytically regular partition $\mathcal{E}$ of W has positive entropy, then the system is not analytically integrable on W.

This criterion can be added to those listed at the beginning of §4.15, p.353 and to the other criteria, also quite remarkable, that emerge from the problems at the end of this section, see problems 13-20. We now quote, without proof, some results on entropy theory and non integrable systems showing that in fact the previously stated non integrability criteria (i), (iii), (iv), and (v) are not empty of content [(ii) has already been discussed in §4.13, Observation (3) Proposition 27, p.336], i.e. the propositions below illustrate other properties of entropy (Proposition 37) or they show that there actually are systems whose non integrability could be decided on the basis of the above criteria (Proposition 38).

37 Proposition. *Let* $\mathbf{a}=(a_i)_{i\in\mathcal{Z}_+0}$, $a_i=0,1,\ldots,p-1$ *be an ergodic sequence.*
(i) The entropy of $\mathbf{a}$ *can be computed as*

$$S(\mathbf{a})=\lim_{N\to\infty}-\frac{1}{N}\sum_{\alpha_0,\ldots,\alpha_{N-1}}\mathbf{p}\left(\begin{smallmatrix}0\ldots N-1\\ \alpha_0\ldots\alpha_{N-1}\end{smallmatrix}\Big|\mathbf{a}\right)\log\mathbf{p}\left(\begin{smallmatrix}0\ldots N-1\\ \alpha_0\ldots\alpha_{N-1}\end{smallmatrix}\Big|\mathbf{a}\right). \tag{4.15.26}$$

(ii) Given $\varepsilon>0$, *there exists* N_ε *such that* $\forall N\ge N_\varepsilon$ *the* p^N *strings* $\left(\begin{smallmatrix}0\ldots N-1\\ \alpha_0\ldots\alpha_{N-1}\end{smallmatrix}\right)$ *of history with length* N, *a priori possible, can be divided into classes* $\mathcal{C}^1_\varepsilon(N)$ *and* $\mathcal{C}^{rare}_\varepsilon(N)$ *such that*

$$\sum_{\alpha_0,\ldots,\alpha_{N-1}\in\mathcal{C}^{rare}_\varepsilon}\mathbf{p}\left(\begin{matrix}0\ \ldots\ N-1\\ \alpha_0\ldots\alpha_{N-1}\end{matrix}\Big|\mathbf{a}\right)<\varepsilon \tag{4.15.27}$$

and for every $\left(\begin{smallmatrix}0\ldots N-1\\ \alpha_0\ldots\alpha_{N-1}\end{smallmatrix}\right)\in\mathcal{C}^1_\varepsilon(N)$

$$e^{-(S(\mathbf{a})+\varepsilon)N}\le\mathbf{p}\left(\begin{matrix}0\ \ldots\ N-1\\ \alpha_0\ldots\alpha_{N-1}\end{matrix}\Big|\mathbf{a}\right)\le e^{-(S(\mathbf{a})-\varepsilon)N} \tag{4.15.28}$$

(iii) The number of elements in $\mathcal{C}^1_\varepsilon(N)$ is such that

$$e^{-(S(\mathbf{a})-\varepsilon)N} \le |\mathcal{C}^1_\varepsilon(N)| \le e^{-(S(\mathbf{a})+\varepsilon)N} \tag{4.15.29}$$

Observations.
(1) This is the "Shannon-McMillan theorem", [25].
(2) Equation (4.15.26) is very useful because it sometimes allows the explicit calculation of $S(\mathbf{a})$. The statement (ii) tells us that if N is large the number of strings of $\mathbf{a}$ that are "really important" is measured by $S(\mathbf{a})$. Furthermore, such strings have about the same probability of appearance, and their number is therefore estimated by Eq. (4.15.29).
In other words, one can think that in a rough (and weak) sense, see Eqs. (4.15.27) and (4.15.28), $\mathbf{a}$ consists of strings of large length each appearing "almost" equally probable (i.e., "almost" equally often) in $\mathbf{a}$.If $\mathbf{a}$ is not ergodic, this last statement is not generally true: this is one of the reasons why the ergodic sequences are interesting.

The following proposition (Hopf-Anosov-Sinai theorem, see [4]) gives an example of an analytic Hamiltonian system which is not analytically integrable.

38 Proposition. *Let $\Sigma \subset \mathcal{R}^d$ be an analytic surface, bounded and with negative curvature. The geodesic motion on Σ (i.e., the motion of a unit mass ideally constrained to Σ) is not analytically integrable because for every analytically regular partition $\mathcal{E}$ of its phase space there exists a dense set of data whose $(\mathcal{E}, \mathbf{t})$ history, $\mathbf{t} = (jt_1)_{j\in\mathcal{Z}_+}$ is mixing and also has positive entropy.*

These last two theorems are two important examples of "ergodic theory" problems. This is a young theory; nevertheless, it is already rich in interesting results and, even more, interesting open problems.

4.15.1 Exercises and Problems

Can one build sequences of preassigned distribution? See Problems 1-12 below.

1. Find examples of sequences $\mathbf{a}$ of symbols $a_i = \pm 1$ with non definite frequencies (*Hint:* For instance 10 symbols 1 followed by 10^{2^1} symbols -1, followed by 10^{2^2} symbols $+1$, etc.)

2. Consider the sequence of symbols $a_i = \pm 1$:

$$a = (1, -1, 1, 1, -1, -1, 1, 1, 1, -1, -1, -1, \ldots).$$

Show that it has well-defined frequencies and that $\mathbf{p}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\middle|\mathbf{a}\right) = \frac{1}{2}$, $\mathbf{p}\left(\begin{smallmatrix}0\,0\\1\,1\end{smallmatrix}\middle|\mathbf{a}\right) = \frac{1}{2}$.

3. Show that the sequence in Problem 2 is non ergodic (*Hint:* Show that Eq. (4.14.35) is false for $j_1 = 0, i_1 = j, \alpha_1 = \beta_1 = 0$.)

4. Find an example of a subset $A \subset \mathcal{T}^\ell$ such that setting $E_0 = A, E_1 = \mathcal{T}^\ell/A$, there is in $\mathcal{T}^\ell$ a point p whose history on the partition $\mathcal{E} = (E_0, E_1)$ with respect to the rotation

$\boldsymbol{\varphi} \to \boldsymbol{\varphi} + \boldsymbol{\omega} \mod 2\pi$, supposed irrational, does not have well-defined frequencies. (*Hint:* Let $\mathbf{a}$ be a sequence of 0's and 1's without well-defined frequencies, see Problem 1; then given p, let $A = \cup_{k, a_k = 0}(p + k\boldsymbol{\omega})$.)

5. Using Proposition 28, §4.13, p.339, and the method of proof of Proposition 30, §4.14, p.342, show that if $E \subset \mathcal{T}^\ell$ is Riemann measurable, then every point of $\mathcal{T}^\ell$ evolving under an irrational rotation transformation visits E with well-defined frequencies.

6. Let $\mathcal{E} = \{0,1\}$, $p_0 = \frac{1}{2}$, $p_1 = \frac{1}{2}$, and consider the probability distributions $(\mathcal{E}, \mathbf{p})$ and $(\mathcal{E}, \mathbf{p})^N$, see Definition 20, §2.23, p.118. Let $A_N(0) \subset \mathcal{E}^N$ be the sequences $\alpha_0, \ldots, \alpha_{N-1}$, $\alpha_j = 0, 1$, in which the symbol 0 appears with frequency closer to $\frac{1}{2}$ than $N^{-\frac{1}{8}}$, i.e.

$$A_N(0) = \Big\{\alpha_0, \ldots, \alpha_{N-1} \,\Big|\, \frac{1}{N}\Big(\sum_{j=0}^{N-1}(1-\alpha_j)\Big) - \frac{1}{2}\Big| < \frac{1}{N^{\frac{1}{8}}}\Big\}.$$

Show that the probability of $A_N(0)$ in $(\mathcal{E}, \mathbf{p})^N$ is such that $\mathbf{p}(A_N(0)) > 1 - \frac{1}{8N^{\frac{3}{4}}}$. (*Hint:* Use Chebysčev's inequality, Proposition 34, p.119; see, also, Proposition 33, p.119.)

7. In the context of Problem 6, regard $A_N(0)$ as a subset $\widetilde{A}_N(0)$ of the space of the infinite sequences $\mathbf{a} = (\alpha_0, \alpha_1, \ldots)$ of 0's and 1's defined by $\mathbf{a} \in \widetilde{A}_N(0) \longleftrightarrow (a_0, \ldots, a_{N-1}) \in A_N(0)$. Show that the sets $\widetilde{A}_{k^2}(0)$ have the finite intersection property, i.e., $\cap_{k=1}^q \widetilde{A}_{k^2}(0) \neq \emptyset, \forall\, q \geq 1$ (*Hint:* Use Problem (6) to note that if $\widetilde{A}_1, \widetilde{A}_2, \widetilde{A}_4, \ldots \widetilde{A}_{k^2}$ are all regarded as subsets in $\mathcal{E}^{k^2}$ in a natural way, they have a probability in $\mathcal{E}^{k^2}$: $\mathbf{p}(A_{k^2}(0)) > 1 - \frac{1}{8k^{\frac{3}{2}}}$. Hence, the complement of the intersection of any number of the A_{k^2}'s has a probability such that

$$\mathbf{p}((\cap A_{k^2}(0))^c) \leq \sum_{k=0}^{\infty} \mathbf{p}(A_{k^2}(0)^c) \leq \frac{1}{8}\sum_{k=1}^{\infty} \frac{1}{k^{\frac{3}{2}}} < 1$$

since $(\cap E_\alpha)^c \subset \cup E_\alpha^c$, in general. Hence, $\cap A_{k^2}(0)$ cannot be empty.)

8. Extend Problem 6 to show that for every given string $(\sigma_1, \ldots, \sigma_s)$ or 0's and 1's, the set $A_N(\sigma_1, \ldots, \sigma_s) \subset (\mathcal{E}, \mathbf{p})^N$ consisting of the strings $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{N-1}) \in \mathcal{E}^N$ in which the string $(\sigma_1, \ldots, \sigma_s)$ appears somewhere, with a frequency differing from 2^{-s} by at most $N^{-\frac{1}{8}}$ is such that

$$\mathbf{p}(A_N(\sigma_1, \ldots, \sigma_s)) > 1 - \frac{\varepsilon_s}{N^{\frac{3}{4}}}$$

for some ε_s. (*Hint:* Proceed as in Problem 6, observing that $A_N(\sigma_1, \ldots, \sigma_s)$ is the set $\Big\{\alpha_0, \ldots, \alpha_{N-1} \mid \frac{1}{N}\Big(\sum_{j=0}^{N-1}(\alpha_j - \sigma_1)^2(\alpha_{j+1} - \sigma_2)^2 \ldots (\alpha_{j+s-1} - \sigma_s)^2\Big) - \frac{1}{2^s} < \frac{1}{N^{1/8}}\Big\}$).

9. Extend Problem 7 as follows: regard $A_N(\sigma_1, \ldots, \sigma_s)$ as a subset $\widetilde{A}_N(\sigma_1, \ldots, \sigma_s)$ in the space of the infinite sequences $\mathbf{a}$ of 0's and 1's defined by $\mathbf{a} \in \widetilde{A}_N(\sigma_1, \ldots, \sigma_s) \longleftrightarrow (a_0, \ldots a_{N-1}) \in A_N(\sigma_1, \ldots, \sigma_s)$. Show that $\exists N_s$, $s = 1, 2, \ldots$, such that $\forall\, n, q \geq 1$:

$$B_{n,q} \stackrel{def}{=} \cap_{s=1}^n \cap_{k=1}^q \cap_{\sigma_0 \ldots \sigma_s}^{0,1} \widetilde{A}_{N_{s+k^2}}(\sigma_0 \ldots \sigma_s) \neq \emptyset.$$

(*Hint:* See the hint to Problem 7 to estimate the probability or the complement of B. One now finds the condition: $\sum_{k=1}^{\infty}\sum_{s=1}^{\infty} \frac{2^s \varepsilon_s}{(N_{s+k^2})^{3/4}} < 1$.)

10. In the context of Problem 9, show that if $\cap_{n,q} B_{n,q} \neq \emptyset$ and $\mathbf{a} \in \cap_{n,q} B_{n,q}$, then $\mathbf{a}$ has well-defined frequencies and:

(i) $\mathbf{p}\left(\begin{smallmatrix} 0 \ldots N-1 \\ \alpha_0 \ldots \alpha_{N-1} \end{smallmatrix} \middle| \mathbf{a}\right) = 2^{-N}, \qquad \forall\, N, \forall\, \alpha_1 \ldots \alpha_N;$

(ii) $\mathbf{a}$ is ergodic and mixing;

(iii) $S_{abs}(\mathbf{a}) = \log 2,\ S(\mathbf{a}) = \log 2$.
(*Hint:* For (ii), check the mixing directly; for (iii), apply, with patience, Definition 21, p.354.)

11. Show that $\cap_{n,q} B_{n,q}$ in Problem 9 is nonempty. (*Hint:* Enumerate, from 1 to ∞, the sets $A_{N_a+k^2}(\sigma_1 \ldots \sigma_s)$ and denote them as $D_1, D_2, \ldots$. Then, by Problem 9, $\cap_{j=1}^m D_j \neq \emptyset$, $\forall\, m$. Let $\mathbf{a}_m \in \cap_{j=1}^m D_j$. Since the sequences $\mathbf{a}_q$ have only two possible entries at each site, there must exist a subsequence $\mathbf{a}_{q_i}$, $q_i \to +\infty$, and a $\mathbf{a}_\infty$ such that $\mathbf{a}_{q_i}$ eventually coincides with $\mathbf{a}_\infty$ on any finite number of sites: $\mathbf{a}_\infty \in \cap_j D_j$.)

12. Extend Problems 6-11 to the case $\mathcal{E} = \{\{0,1\}, p_0 > 0, p_1 > 0, p_0 + p_1 = 1, p_0 \neq \frac{1}{2}$. Show that there are sequences of 0's and 1's such that $S_{abs}(\mathbf{a}) = 1og2$, $S(\mathbf{a}) = -p_0 \log p_0 - p_1 \log p_1 < A_{abs}(\mathbf{a})$.

Other necessary integrability criteria emerge from the following series of problems together with other remarkable properties of integrable systems.

13. Let $A_1, \ldots, A_\ell$ be ℓ prime integrals for an ℓ-degrees-of-freedom Hamiltonian system on $W \subset \mathcal{R}^{2\ell}$, or $W \subset \mathcal{R}^\ell \times \mathcal{T}^\ell$ or $W \subset \mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2})$, $\ell_1 + \ell_2 = \ell$, open. Call $\mathbf{A}(W)$ the set of the values of $(A_1, \ldots, A_\ell)$ on W: $\mathbf{A}(W) \subset \mathcal{R}^\ell$. Suppose that the equation $\mathbf{A}(\mathbf{p}, \mathbf{q}) = \mathbf{a}$ can be inverted with nonzero Jacobian near $\mathbf{p}_0, \mathbf{q}_0, \mathbf{a}_0$ $\mathbf{p} = \boldsymbol{\alpha}(\mathbf{a}, \mathbf{q})$ so that $\mathbf{A}(\boldsymbol{\alpha}(\mathbf{a}, \mathbf{q}), \mathbf{q}) = \mathbf{a}$. Define the $\ell \times \ell$ matrices:

$$M_{ij} = \frac{\partial A_i}{\partial p_j}, \quad N_{ij} = \frac{\partial A_i}{\partial q_j}, \quad T_{ij} = \frac{\partial \alpha_i}{\partial q_j}, \quad R_{ij} = \frac{\partial \alpha_i}{\partial a_j}, .$$

Study the "Hamilton-Jacobi" equations:

$$\mathbf{A}(\frac{\partial s}{\partial \mathbf{q}}, \mathbf{q}) = \mathbf{a}, \qquad \text{i.e.} \qquad \frac{\partial s}{\partial \mathbf{q}} = \boldsymbol{\alpha}(\mathbf{a}, \mathbf{q})$$

and find conditions "guaranteeing their solubility" *near* $\mathbf{q}_0, \mathbf{a}_0$. Check that conditions could be $\{A_i, A_j\} = 0$, $\forall\ i, j = 1, \ldots, \ell$, i.e.,

$$\sum_{s=1}^{\ell} \left(\frac{\partial A_i}{\partial p_s} \frac{\partial A_j}{\partial q_s} - \frac{\partial A_i}{\partial q_s} \frac{\partial A_j}{\partial p_s} \right)$$

(see, also, Definition 19, §3.12). (*Hint:* It is only needed that the differential form $\boldsymbol{\alpha} \cdot d\mathbf{q}$ be exact i.e., $\frac{\partial \alpha_i}{\partial q_j} = \frac{\partial \alpha_j}{\partial q_i}$ or $T_{ij} = T_{ji}$. By the implicit function theorem and by the chain differentiation rule, it follows from $\mathbf{A}(\boldsymbol{\alpha}(\mathbf{a}, \mathbf{q}), \mathbf{q}) \equiv \mathbf{a}$ that

$$\sum_{s=1}^{\ell} \frac{\partial A_i}{\partial p_s} \frac{\partial \alpha_s}{\partial a_j} = \delta_{ij}, \qquad \text{and} \ \sum_{s=1}^{\ell} \left(\frac{\partial A_i}{\partial p_s} \frac{\partial \alpha_s}{\partial q_j} \right) + \frac{\partial A_i}{\partial q_j} = 0;$$

i.e., with the above notations, $NR = 1$ and $MT + N = 0$. So, since $T = -M^{-1}N$, the integrability condition becomes

$$M^{-1}N = (M^{-1}N)^T = N^T (M^{-1})^T \longleftrightarrow NM^T = MN^T$$

because the Jacobian determinant $\det M \neq 0$. The last expression once written explicitly, yields the result ("Liouville's theorem").)

14. Show the following properties of the Poisson bracket see Definition 19, §3.12:

$$\{F, G\} = -\{G, F\},$$
$$\{F, GL\} = \{F, G\}L + \{F, L\}G,$$
$$\{F, \{G, L\}\} + \{G, \{L, F\}\} + \{L, \{F, G\}\} = 0.$$

Two observables on phase space F, G are said to be "in involution" if $\{F, G\} = 0$.

15. In the context of Problems 13 and 14, suppose that $A_1, \ldots, A_\ell$ are ℓ prime integrals in involution. Consider the completely canonical transformation $\mathcal{C}$ generated by the function $(\mathbf{a}, \mathbf{q}) \to s(\mathbf{a}, \mathbf{q})$ in Problem 13 (via $\boldsymbol{\kappa} = \frac{\partial s(\mathbf{a},\mathbf{q})}{\partial \mathbf{a}}$, $\mathbf{p} = \frac{\partial s(\mathbf{a},\mathbf{q})}{\partial \mathbf{q}}$. Denote it $(\mathbf{a}, \boldsymbol{\kappa}) = \mathcal{C}(\mathbf{p}, \mathbf{q})$. Show that $H(\mathcal{C}^{-1}(\mathbf{a}, \boldsymbol{\kappa})) = h(\mathbf{a})$ is $\boldsymbol{\kappa}$ independent. (*Hint:* Since the $\mathbf{A}$'s are prime integrals ($\mathbf{A} = \mathbf{a}(\mathbf{p}, \mathbf{q})$) and the map $(\mathbf{p}, \mathbf{q}) \longleftrightarrow (\mathbf{a}, \boldsymbol{\kappa})$ is completely canonical, it must be that

$$\dot{\mathbf{a}} = \frac{\partial H(\mathcal{C}^{-1}(\mathbf{a}, \boldsymbol{\kappa}))}{\partial \boldsymbol{\kappa}} = \mathbf{0}.$$

i.e., $H(\mathcal{C}^{-1}(\mathbf{a}, \boldsymbol{\kappa}))$ is $\boldsymbol{\kappa}$ independent.)

16. Using the fact that the completely canonical transformations preserve the Poisson brackets, see Observation (2), p.237, to Corollary 25, §3.12, show that a necessary condition for the canonical integrability of a Hamiltonian system on a region W of phase space is the existence in W of ℓ independent prime integrals in involution.

17. Show that a necessary and sufficient condition in order that $A \in C^\infty(W)$ be a prime integral for a regular Hamiltonian system on W is that $\{A, H\} = 0$, if H is the Hamiltonian function. More generally, if $S_t(\mathbf{p}, \mathbf{q})$, $t \in J$, denotes a solution to the Hamilton equations in W and $F \in C^\infty(W)$, show that

$$\frac{d}{dt} F(S_t(\mathbf{p}, \mathbf{q})) = \{H, F\}(S_t(\mathbf{p}, \mathbf{q})), \qquad \forall\, t \in J.$$

Here $W \subset \mathcal{R}^{2\ell}$ or $\mathcal{R}^\ell \times \mathcal{T}^\ell$ or $\mathcal{R}^\ell \times (\mathcal{R}^{\ell_1} \times \mathcal{T}^{\ell_2})$, $\ell_1 + \ell_2 = \ell$ is open. (*Hint:* Just compute the derivative of F using the Hamilton equations, to express $\dot{\mathbf{p}}, \dot{\mathbf{q}}$, and the definition of the Poisson bracket.)

18. Let W be as in the above problems and let $H \in C^\infty(W)$ be a regular Hamiltonian function. Assume that H is integrable on W and let I be the integrating transformation $I : W \longleftrightarrow V \times \mathcal{T}^\ell$, $V \subset \mathcal{R}^\ell$, let $(\mathbf{A}, \boldsymbol{\varphi}) = I(\mathbf{p}, \mathbf{q})$ and denote $\boldsymbol{\omega}(\mathbf{A})$ the pulsations of the quasi-periodic motions on the torus $\{\mathbf{A}\} \times \mathcal{T}^\ell$. We say that the system is "non isochronous" in W if the matrix $J_{ij}(\mathbf{A}) = \frac{\partial \omega_i}{\partial A_j}(\mathbf{A})$ has a non vanishing determinant.
Show that any prime integral $B \in C^\infty(W)$ for a non isochronous integrable Hamiltonian system must be a function of $A_1, \ldots, A_\ell$ introduced above. (*Hint:* Let $B = b(\mathbf{A}, \boldsymbol{\varphi})$ be a prime integral in the $(\mathbf{A}, \boldsymbol{\varphi})$ variables. It must be $b(\mathbf{A}, \boldsymbol{\varphi}) \equiv b(\mathbf{A}, \boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{A})t), \forall\, t \in \mathcal{R}$. If the components of $\boldsymbol{\omega}(\mathbf{A})$ are rationally independent the points $\boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{A})t$, $t \in \mathcal{R}$ densely cover $\mathcal{T}^\ell$; hence, for such $\mathbf{A}$'s, B must depend only on $\mathbf{A}$ and not on $\boldsymbol{\varphi}$. However, if $\det J \neq 0$, the set of $\mathbf{A}$'s in V such that $\boldsymbol{\omega}(\mathbf{A})$ has rationally independent coordinates is dense in V (see Problems 9 and 15, §5.10, p.477 and 478). Hence, B must always depend only on $\mathbf{A}$.)

19. There is a theorem by Arnold concerning the case when W is an invariant open bounded set for a regular Hamiltonian flow generated by $H \in C^\infty(W)$ and on W one can define ℓ independent prime integrals $\mathbf{A} = (A_1, \ldots, A_\ell)$ in involution (see Problem 14), with $A_1 \equiv H$ and such that the sets $\mathbf{A}(\mathbf{p}, \mathbf{q}) = \mathbf{a}$ are, for $\mathbf{a} \in \mathbf{A}^1(W)$, regular closed bounded and connected surfaces in W. Then H is integrable on W.
Are there systems integrable but not canonically integrable? (*Answer:* If $\ell = 2$, some partial results are known ([43]; another partial answer is in Problem 22 below). The proof of Arnold's theorem can be found on page 269 of [1]).

20. Find an example of a Hamiltonian system whose motions are all quasi-periodic but which is not integrable. (*Hint:* consider two point masses free on a circle and on a line, respectively; let their positions be determined by $(\varphi_1, q_2) \in \mathcal{T}^1 \times \mathcal{R}$ if φ_1 is the angular position of the first particle and q_2 the position of the second. Let $H(p_1, p_2, \varphi_1, q_2) = \frac{p_1^2}{2}$.)

21. Suppose that in the region W an analytic Hamiltonian $H(\mathbf{p},\mathbf{q})$ admits ℓ independent prime integrals $\mathbf{A}=(A_1,\ldots,A_\ell)$ and $H=A_1$. Suppose that the surfaces $\mathbf{A}=\mathbf{a}$ are tori of dimension ℓ. Write their parametric equations as $\mathbf{p}=\mathbf{P}(\mathbf{a},\boldsymbol{\varphi}),\mathbf{Q}(\mathbf{a},\boldsymbol{\varphi})$ and suppose that the evolution is $\boldsymbol{\varphi}\to\boldsymbol{\varphi}+\boldsymbol{\omega}(\mathbf{A})\,t$: i.e. suppose that all motions are quasi periodic. If $\det\frac{\partial\boldsymbol{\omega}(\mathbf{A})}{\partial\mathbf{A}}\neq 0$, i.e. if the system is anisochronous, then $\{A_i,A_j\}=0,\forall\, i,j=1,\ldots,\ell$. (*Hint:* Suppose that $\{A_i,A_j\}\neq 0$ then evolve an initial datum $(\mathbf{a},\boldsymbol{\varphi})$ with the Hamilton equations with Hamiltonian A_i for a small time ε and then with the Hamiltonian $A_1=H$ for a long time t. Since A_i and H have zero Poisson bracket the two evolutions "commute" and the final datum has to be the same as the one obtained by first evolving $(\mathbf{a},\boldsymbol{\varphi})$ for a time t with the Hamilton equations for H and then for a time ε with A_j. In the first case the result will be a datum $(\mathbf{a}',\boldsymbol{\varphi}'+\boldsymbol{\omega}'t)$ with $\mathbf{a}',\boldsymbol{\varphi}',\boldsymbol{\omega}'$ close $O(\varepsilon)$ to $(\mathbf{a},\boldsymbol{\varphi},\boldsymbol{\omega})$; in the second case the result will be $(\mathbf{a}'',\boldsymbol{\varphi}''+\boldsymbol{\omega}t)$ with $\mathbf{a}'',\boldsymbol{\varphi}''$ close to $\mathbf{a},\boldsymbol{\varphi}$ within $O(\varepsilon)$. *However* since $\{A_i,A_j\}\neq 0$ *it is* $\boldsymbol{\omega}'\neq\boldsymbol{\omega}$ and this is a contradiction for t large.)

22. Check that combining Problems 21 and 19 above a new criterion of completely canonical integrability follows.

23. Let $k\to f(k)$ be a function defined for $k=1,2,\ldots$ such that $0\le f(k)$, $f(k+h)\le f(k)+f(h)$, for all $h,k=1,2,\ldots$. Show that

$$\lim_{k\to+\infty}\frac{f(k)}{k}=\inf_k\frac{f(k)}{k}\stackrel{def}{=}s$$

and apply this result to prove the existence of the limit (4.15.2) by showing that $f(k)=\log N_{abs}(k,\mathbf{a})$ has the above subadditivity properties. (*Hint:* Let $\varepsilon>0$ and let k_ε be such that $s\le\frac{1}{k_\varepsilon}f(k_\varepsilon)\le s+\varepsilon$; write $k=hk_\varepsilon+p$ with $h=0,1,\ldots$ and $p=0,1,\ldots,k_\varepsilon-1$ and note that $s\le k^{-1}f(k)\le(hk_\varepsilon+p)^{-1}(hf(k_\varepsilon)+f(p))\xrightarrow[k\to+\infty]{}k_\varepsilon^{-1}f(k_\varepsilon)<s+\varepsilon$.)

5

Stability Properties for Dissipative and Conservative Systems

5.1 A Mathematical Model for the Illustration of Some Properties of Dissipative Systems

In various possible senses, the stability properties of motions are more easily analyzed in systems moving in the presence of friction, as already noted in Chapter 2.

Therefore, we shall mainly concentrate our attention on such systems, studying some stability questions selected among others because they seem particularly significant for the generality of the methods used to treat them.

Similar questions will later be asked about conservative systems. However, the answers, when known, will be much harder to obtain.

The gyroscope is, in some sense, the prototype for systems with many degrees of freedom. In fact, general systems of linear oscillators trivially reduce to systems of independent one dimensional oscillators, as explained §4.1-4.4 in the conservative cases; this remains true even in the presence of linear friction.

On the other hand, the gyroscope with friction, or even some of its particular cases, already presents many of the possibilities and difficulties that can be met in more complex systems.

For this reason, in the upcoming sections, we shall illustrate the general theory through the treatment of a single example, described below and drawn from the gyroscope theory, which will be used to motivate the successive steps of a theory and of a method of analysis which, as will become evident, is applicable to many other dissipative systems as well.

The example is given by Eqs. (5.1.18) and (5.1.19) and this section is devoted to their gyroscopic interpretation.

We consider a rigid body consisting of N masses, $m_1, m_2, ..., m_N > 0$, with a fixed point O (all the constraints being ideal) immersed in a viscous fluid opposing to the motion a frictional force at the i-th point:

$$-\lambda m_i \dot{\mathbf{x}}^{(i)} \tag{5.1.1}$$

The moment of the frictional force with respect to 0 is then given by

$$-\lambda \sum_{i=1}^{N} m_i (P_i - O) \wedge \dot{\mathbf{x}}^{(i)} = -\lambda \sum_{i=1}^{N} m_i (P_i - O) \wedge (\boldsymbol{\omega} \wedge (P_i - O)) = \lambda I \boldsymbol{\omega} \tag{5.1.2}$$

with the notations of §4.11.

The second cardinal equation,[1] implies

$$I\dot{\boldsymbol{\omega}} = -\lambda I \boldsymbol{\omega} - \boldsymbol{\omega} \wedge I\boldsymbol{\omega}, \tag{5.1.3}$$

where $\boldsymbol{\omega}$ is the vector whose components in a co-moving frame $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ are the derivatives of the corresponding components of $\boldsymbol{\omega}$ in the same frame. Equation (5.1.3) extends Eq. (4.11.31) to the case when the moment of the external forces is -$\lambda I \boldsymbol{\omega}$ instead of $\mathbf{0}$.

Assume that the co-moving frame has been fixed once and for all so that the inertia matrix I is diagonal, see Eq. (4.11.9), p.308, with elements

$$I_1, I_2, I_3. \tag{5.1.4}$$

In order to obtain nontrivial motions, it will be convenient to imagine that the system is subject to the action of other forces having a moment $\mathbf{M}$ with respect to O. Otherwise, as is intuitively clear and as we shall shortly see, the system will just stop. The simplest force laws are those with moment $\mathbf{M}$ having constant components on the axes of $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$:

$$\mathbf{M} = R_1 \mathbf{i}_1 + R_2 \mathbf{i}_2 + R_3 \mathbf{i}_3 \tag{5.1.5}$$

or those with moment components in $(O; \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$ dependent only upon the angular velocity

$$\mathbf{M}'(\boldsymbol{\omega}) = R_1'(\boldsymbol{\omega})\mathbf{i}_1 + R'(\boldsymbol{\omega})_2(\boldsymbol{\omega})\mathbf{i}_2 + R_3'(\boldsymbol{\omega})\mathbf{i}_3 \tag{5.1.6}$$

which can be imagined (as in the examples below) generated by some “inner mechanisms” regulating their action as a function of the motion of the body.

In the presence of forces with moments of Eqs. (5.1.5) and (5.1.6) added to the friction forces, the equations of motion of the system would become

$$I\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \wedge I\boldsymbol{\omega} - \lambda I \boldsymbol{\omega} + \mathbf{M} + \mathbf{M}'(\boldsymbol{\omega}). \tag{5.1.7}$$

Even in the simplest situations, e.g.if

[1] $\dot{\mathbf{K}}_O = -\lambda I \boldsymbol{\omega}$.

$$\mathbf{M} = R\,\mathbf{i}_3, \qquad \mathbf{M}'(\boldsymbol{\omega}) = \text{ linear function of } \boldsymbol{\omega} \tag{5.1.8}$$

it could a priori happen that the differential equation E. (5.1.7) admits solutions $t \to S_t(\overline{\boldsymbol{\omega}})$, with suitable initial datum $\overline{\boldsymbol{\omega}}$, diverging as $t \to +\infty$.[2]

We wish to avoid having to deal with such phenomena, too idealized from a physical point of view, since it is clear that any real system "breaks down into pieces" if $\boldsymbol{\omega}$ reaches too large a value, when the centrifugal forces exceed the materials resistance. This is done by supposing that the friction coefficient λ has some extra dependence on $\boldsymbol{\omega}$. For instance,

$$\lambda(\boldsymbol{\omega}) = (\lambda_1 + \lambda_2 \boldsymbol{\omega}^2), \quad \lambda(\boldsymbol{\omega}) = (\lambda_1 + \lambda_2' \omega_1^2 + \lambda_2'' \omega_2^2 + \lambda'''{}_2 \omega_3^2) \tag{5.1.9}$$

which is a special case of the more general and realistic friction model in which Eq. (5.1.1) is replaced by $-\lambda\mu_i \left(1 + (\dot{\mathbf{x}}^{(i)} \cdot L_i \dot{\mathbf{x}}^{(i)})\right) \dot{\mathbf{x}}^{(i)}$, $\lambda_i, \mu_i > 0$ and L_i are 3×3 positive-definite matrices.

Summarizing the above discussion, the mechanical system whose properties we wish to analyze will be described by the equation

$$I\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \wedge I\boldsymbol{\omega} - \lambda(\boldsymbol{\omega}) I\boldsymbol{\omega} + \mathbf{M} + \mathbf{M}'(\boldsymbol{\omega}), \tag{5.1.10}$$

where $\lambda(\boldsymbol{\omega})$ is given by Eq. (5.1.9) and $\mathbf{M}'(\boldsymbol{\omega})$ is a linear function of $\boldsymbol{\omega}$.

The above system is general enough to present a great variety of phenomena. For simplicity, we shall impose further restrictions, studying the following particular case of Eq. (5.1.10).

(i) The rigid body is a gyroscope: $I_1 = I_2, \quad I_3 = J.$ (5.1.11)

(ii) $\mathbf{M} = R\,\mathbf{i}_3,\ R > 0.$ (5.1.12)

(iii) $\mathbf{M}'(\boldsymbol{\omega}) = \alpha_1 \omega_1 \mathbf{i}_1 + \alpha_2 \omega_2 \mathbf{i}_2, \quad \alpha_1 = \alpha_2 = \alpha > 0.$ (5.1.13)

(iv) $\lambda(\boldsymbol{\omega})$ is given by the first or second of Eqs. (5.1.9),(5.1.13).

It might be useful to have in mind a physical representation of the special system mathematically described by Eqs. (5.1.9)-(5.1.13): think of the body as consisting of six masses m located at the points $\pm\varrho\,\mathbf{i}_1, \pm\varrho\,\mathbf{i}_2, \pm\varrho'\,\mathbf{i}_3$. Then

$$I = I_1 = I_2 = 2m(\varrho^2 + {\varrho'}^2), \qquad J = I_3 = 4m\varrho^2. \tag{5.1.14}$$

The force given by Eq. (5.1.12) can be imagined to be generated by small "jet motors" located at the four points $\pm\varrho\mathbf{i}_1, \pm\varrho\mathbf{i}_2$ of the $z_3 = 0$ plane, producing a thrust f identical at each site and perpendicular to the coordinate axis on which the site lies and parallel to the $z_3 = 0$ plane. The moment of such forces is

[2] However, in this case, the global existence of the solutions, assuming λ constant and $\mathbf{M}'$ linear in $\omega_1, \omega_2, \omega_3$, follows from an a priori estimate, for $t \in \mathcal{R}_+$. Let $\boldsymbol{\Omega} = I\boldsymbol{\omega}$ and multiply both sides of Eq. (5.1.7) scalarly by $\boldsymbol{\Omega}$. It follows: $\frac{1}{2}\frac{d\boldsymbol{\Omega}^2}{dt} \le K\boldsymbol{\Omega}^2 + K'$ for some $K, K' > 0$, which implies $(K\boldsymbol{\Omega}(t)^2 + K') \le (K\boldsymbol{\Omega}(0)^2 + K')e^{2Kt}, \forall\, t \ge 0$.

$$\mathbf{M} = 4\varrho\, f\, \mathbf{i}_3 \tag{5.1.15}$$

like Eq. (5.1.12) with $R = 4\varrho f$. The other force given by Eq. (5.1.13) is generated by small jet motors located at the two points on the axis $\mathbf{i}_3$, exerting a thrust along $\mathbf{i}_1$ and $\mathbf{i}_2$, respectively, with intensities

$$f'\omega_2\mathbf{i}_1 \quad \text{and} \quad f'\omega_1\mathbf{i}_2 \tag{5.1.16}$$

and, therefore, their moment is

$$f'\varrho'(\omega_1\mathbf{i}_1 + \omega_2\mathbf{i}_2) \tag{5.1.17}$$

like Eq. (5.1.13) with $\alpha = f'\,\varrho'$.

The somewhat bizarre force given by Eq. (5.1.16) must be thought of as generated by jets producing a thrust proportional to the amount of air entering them per unit time, supposing them to be oriented as $\mathbf{i}_2$ and $\mathbf{i}_1$ respectively, and orthogonal to $\mathbf{i}_3$. The amount of air entering the jets per unit time is in this way proportional to ω_2z and ω_1, respectively.

Obviously, if $f' \neq 0$, the gyroscope will tend to increase its rotation speed around the axes $\mathbf{i}_1, \mathbf{i}_2$, but not indefinitely: just as long as the system reaches a rotation speed causing so strong a friction as to compensate for the force of the motor (this is what actually happens if $\lambda_1', \lambda_2^{''}, \lambda_2^{'''} > 0$).

Explicitly writing Eq. (5.1.10) by components, given the assumptions of Eqs. (5.1.9)-(5.1.13), it is

$$\begin{aligned}
\dot{\omega}_1 &= -\left(\lambda_1 + \lambda_2\boldsymbol{\omega}^2\right)\omega_1 + \alpha\,\omega_1 - \omega_2\omega_3 \\
\dot{\omega}_2 &= -\left(\lambda_1 + \lambda_2\boldsymbol{\omega}^2\right)\omega_2 + \alpha\,\omega_2 + \omega_1\omega_3 \\
\dot{\omega}_3 &= -\left(\lambda_1 + \lambda_2\boldsymbol{\omega}^2\right)\omega_3 + R,
\end{aligned} \tag{5.1.18}$$

with $R, \alpha, \lambda_1, \lambda_2 > 0$, if the first of Eqs. (5.1.9) is assumed and if $(J-I)/I = 1$, a case to which one can reduce by the change of variables $\omega_i' = \omega_i \frac{J-I}{I}$; α, R are real numbers supposed positive, for definiteness.

If the second of Eqs. (5.1.9) is assumed, then

$$\begin{aligned}
\dot{\omega}_1 &= -\left(\lambda_1 + \lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2'''\omega_3^2\right)\omega_1 + \alpha\,\omega_1 - \omega_2\omega_3 \\
\dot{\omega}_2 &= -\left(\lambda_1 + \lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2'''\omega_3^2\right)\omega_2 + \alpha\,\omega_2 + \omega_1\omega_3 \\
\dot{\omega}_3 &= -\left(\lambda_1 + \lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2'''\omega_3^2\right)\omega_3 + R,
\end{aligned} \tag{5.1.19}$$

with $R, \alpha, \lambda_1, \lambda_2', \lambda_2'', \lambda_2^{'''} > 0$. Define $\lambda_2 = \min(\lambda_1', \lambda_2''.\lambda_2^{'''}) > 0$.

A symmetry in Eq. (5.1.18), absent in Eq. (5.1.19), leads to the elimination of one of the variables. In fact, if $\omega^2 \stackrel{def}{=} \omega_1^2 + \omega_2^2$, we find, by multiplying the first of Eqs. (5.1.18) by ω_1, and the second by ω_2 and adding them:

$$\begin{aligned} \frac{1}{2}\frac{d\omega^2}{dt} &= -\,(\lambda_1 + \lambda_2\omega^2 + \lambda_2\omega_3^2)\omega^2 + \alpha\omega^2, \\ \frac{d\omega_3^2}{dt} &= -\,(\lambda_1 + \lambda_2\boldsymbol{\omega}^2)\omega_3 + R, \end{aligned} \tag{5.1.20}$$

with $\lambda_1, \lambda_2, R, \alpha > 0$: much simpler as it involves only two unknowns, ω^2, ω_3..

5.2 Stationary Motions for a Dissipative Gyroscope

Remark that Eqs. (5.1.18) and (5.1.19) admit global solutions in the future.

1 Proposition. *Equation (5.1.19) admits a solution $t \to S_t(\boldsymbol{\omega}_0)$, $t \in \mathcal{R}_+$, for every initial datum $\boldsymbol{\omega}_0 \in \mathcal{R}^3$.*
Furthermore, if $\lambda_2 = \min(\lambda_1', \lambda_2''.\lambda_2''')$ and $\Omega = (\frac{2R}{\lambda_2})^{\frac{1}{3}} + (\frac{2|\alpha=\lambda_1|}{\lambda_2})^{\frac{1}{2}}$:

$$\text{(i)} \qquad |S_t(\boldsymbol{\omega}_0)| \le |\boldsymbol{\omega}_0| + \Omega, \quad \forall\, t \ge 0 \tag{5.2.1}$$

$$\text{(ii)} \quad |S_t(\boldsymbol{\omega}_0)| \le 2\Omega, \quad \forall\, t \ge \frac{(|\boldsymbol{\omega}_0|^2 - 4\Omega^2)}{2\lambda_2\Omega^4}. \tag{5.2.2}$$

Observations.
(1) Equation (5.2.1) means that the trajectory of the motions of the $\boldsymbol{\omega}$'s are bounded uniformly for $t \ge 0$.
(2) Equation (5.2.2) means that all motions take place inside the ball with radius 2Ω after a finite transient time (which may depend upon the initial datum).

Proof. To show global existence, it suffices to show, on the basis of Definition 3 and Proposition 5, §2.5, p.28, an a priori estimate, i.e., it suffices to show that if $t \to S_t(\boldsymbol{\omega}_0)$ is a solution to Eq. (5.1.19) for $t \in [0,T]$ with datum $\boldsymbol{\omega}_0$, then it verifies the inequality (5.2.1), $\forall\, t \in [0,T]$.

This is a simple consequence of the structure of Eq. (5.1.19). In fact, let $\boldsymbol{\omega} = S_t(\boldsymbol{\omega}_0)$ and multiply the equations by $\omega_1, \omega_2, \omega_3$, respectively; adding the results yields

$$\frac{d}{dt}\frac{1}{2}\boldsymbol{\omega}^2 = -\lambda(\boldsymbol{\omega})\boldsymbol{\omega}^2 + \alpha\,(\omega_1^2 + \omega_2^2) + R\,\omega_3 \le |\alpha - \lambda_1|\boldsymbol{\omega}^2 - \lambda_2\boldsymbol{\omega}^4 + |\omega_3|. \tag{5.2.3}$$

Assuming the inequalities

$$\frac{\lambda_2}{2}\boldsymbol{\omega}^4 > |\alpha - \lambda_1|\boldsymbol{\omega}^2, \qquad \frac{\lambda_2}{2}\boldsymbol{\omega}^4 > R|\boldsymbol{\omega}| \tag{5.2.4}$$

the right-hand side of Eq. (5.2.3) is negative. Hence, if initially $|\boldsymbol{\omega}_0| > \Omega$ with

$$\Omega \stackrel{def}{=} \Big(\frac{2R}{\lambda_2}\Big)^{\frac{1}{3}} + \Big(\frac{2|\alpha - \lambda_1|}{\lambda_2}\Big)^{\frac{1}{2}}, \tag{5.2.5}$$

the quantity $|S_t(\boldsymbol{\omega}_0)| \equiv |\boldsymbol{\omega}|$ must decrease as t grows, at least until it becomes $\le \Omega$. This implies both global existence and the estimate (5.2.1).

To find the estimate (5.2.2), note that $|\boldsymbol{\omega}| \geq 2\Omega$ implies that the right hand side of Eq. (5.2.3) is smaller than $-\lambda_2\Omega^4$. Hence, as long as $|S_t(\boldsymbol{\omega}_0)| \geq 2\Omega$, one must have

$$|S_t(\boldsymbol{\omega}_0)|^2 \leq |\boldsymbol{\omega}_0|^2 - 2\lambda_2\Omega^4 t \tag{5.2.6}$$

which means that for $t \geq \frac{|\boldsymbol{\omega}_0|^2-4\Omega^2}{2\lambda_2\Omega^4}$, it will be $|S_t(\boldsymbol{\omega}_0)| \leq 2\Omega$. mbe

In general, the simplest information about the nature of the motions described by a differential equation can be obtained through the study of stationary solutions.

2 Proposition. *Equation (5.1.19) has,* $\forall\, R > 0$ *and* $\forall\, \alpha > 0$*, a unique stationary solution* $\widehat{\boldsymbol{\omega}}$*. This solution has* $\widehat{\omega}_1 = \widehat{\omega}_2 = 0$*, while* ω_3 *is the unique real solution to the equation*

$$-(\lambda_1 + \lambda_2^{'''}\widehat{\omega}_3^2)\widehat{\omega}_3 + R = 0. \tag{5.2.7}$$

Proof. Setting $\dot{\omega}_1 = \dot{\omega}_2 = 0$ in the first two of Eqs. (5.1.19) and imagining[3] known $\lambda(\widehat{\boldsymbol{\omega}})$ and $\widehat{\omega}_3$, one obtains two homogeneous linear equations for $\widehat{\omega}_1, \widehat{\omega}_2$ with determinant

$$(\alpha - \lambda(\widehat{\boldsymbol{\omega}}))^2 + \widehat{\omega}_3^2 \tag{5.2.8}$$

which vanishes only for $\widehat{\omega}_3 = 0$ and $\alpha = \lambda(\widehat{\boldsymbol{\omega}})$, but the third of Eqs. (5.1.19) does not admit a stationary solution with $\widehat{\omega}_3 = 0$. Hence, Eq. (5.2.8) does not vanish and, therefore, $\widehat{\omega}_1 = \widehat{\omega}_2 = 0$, which in turn implies that $\widehat{\omega}_3$ has to verify Eq. (5.2.7). This equation admits just one solution by the strict monotonicity in ω_3 of the left-hand side. mbe

A natural question is: how does the actual motion of the gyroscope look if the angular velocity is $\widehat{\boldsymbol{\omega}}$?

3 Proposition. *The motion of the gyroscope corresponding to the stationary solution* $\widehat{\boldsymbol{\omega}}$ *of Eq. (5.1.19) is a rotation with constant angular velocity* $\widehat{\omega}_3$ *around the axis* $\mathbf{i}_3$*, which remains fixed in space.*

Proof. Let $t \to (\theta(t), \varphi(t), \psi(t))$ be a description of the motion in terms of the Euler angles (θ, φ, ψ) of $(O; \mathbf{i}_1\mathbf{i}_2, \mathbf{i}_3)$ with respect to the fixed reference frame $(O, \mathbf{i}, \mathbf{j}, \mathbf{k})$.

From Eqs. (4.11.12), (4.11.13), and (4.11.14),[4] p.309, one deduces the relationship between $\dot{\theta}(t), \dot{\varphi}(t), \dot{\psi}(t)$, and the vector $\boldsymbol{\omega}(t)$. In general,

$$\dot{\theta} = \omega_1 \cos\psi - \omega_2 \sin\psi, \tag{5.2.9}$$

[3] $\lambda(\boldsymbol{\omega})$ denotes $\lambda_1 + \lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2^{'''}\omega_3^2$.

[4] Without the bars, since now there is no need of them.

$$\dot{\varphi} = \frac{\omega_1 \sin\psi - \omega_2 \cos\psi}{\sin\theta} \tag{5.2.10}$$

$$\dot{\psi} = \omega_3 - \frac{\cos\theta}{\sin\theta}(\omega_1 \sin\psi + \omega_2 \cos\psi). \tag{5.2.11}$$

Letting $\omega_2 = \omega_2 = 0$ and $\omega_3 = \widehat{\omega}_3$, one deduces from Eq. (5.2.9) that θ is constant ($\dot{\theta} = 0$). Hence suppose, without loss of generality, to have fixed $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ so that $\theta(0) \neq 0$ or π.

The second equation, Eq. (5.2.10), implies$\dot{\varphi} = 0$. Hence, φ is a constant. Since θ and φ determine the position of $\mathbf{i}_3$ in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$, it follows that $\mathbf{i}_3$ is fixed in $(O; \mathbf{i}, \mathbf{j}, \mathbf{k})$ and, therefore, the system rotates around $\mathbf{i}_3$ (fixed) with angular velocity given by $\dot{\psi} = \widehat{\omega}_3$, by Eq. (5.2.11). mbe

We can now begin the study of non stationary motions. If $\alpha < \lambda_1$ the motions are particularly simple.

4 Proposition. *If $\alpha < \lambda_1$ the solutions $t \to S_t(\boldsymbol{\omega})$ of Eq. (5.1.19) with initial datum $\boldsymbol{\omega}$ verify*

$$|S_t(\boldsymbol{\omega}) - \widehat{\boldsymbol{\omega}}| < c(|\boldsymbol{\omega}|)\, e^{-(\lambda_1 - \alpha)t}, \tag{5.2.12}$$

where $c(x)$ is a suitable increasing function of $x \in \mathcal{R}_+$.
The corresponding motion of the gyroscope tends asymptotically to become a uniform rotation with angular velocity $\widehat{\omega}_3$ around the axis $\mathbf{i}_3$ which in turn tends to acquire a fixed position in $(O, \mathbf{i}, \mathbf{j}, \mathbf{k})$, the fixed reference frame.

More precisely, if $t \to (\theta(t), \varphi(t), \psi(t))$ is the description of the motion of the Euler angles, whose angular velocity is $\boldsymbol{\omega}(t) = S_t(\boldsymbol{\omega})$, for $t \geq 0$, there exist constants $t_1 > 0, C_1 > 0, \overline{\theta}, \overline{\varphi}, \overline{\psi}$, depending on the initial data and such that

$$\begin{aligned} |\theta(t) - \overline{\theta}| &\leq C_1\, e^{-(\lambda_1 - \alpha)\, t} \\ |\varphi(t) - \overline{\varphi}| &\leq \frac{C_1}{\sin\overline{\theta}}\, e^{-(\lambda_1 - \alpha)\, t} \\ |\psi(t) - \overline{\psi} - \widehat{\omega}_3 t| &\leq \frac{C_1}{\sin\overline{\theta}}\, e^{-(\lambda_1 - \alpha)\, t}. \end{aligned} \tag{5.2.13}$$

For instance, C_1 can be chosen as $C_1 = \frac{4\, c(\Omega + |\boldsymbol{\omega}|)}{\lambda_1 1 - \alpha}$, see Eq. (5.2.12).

Proof. First check that Eq. (5.2.12) implies Eq. (5.2.13). In fact, Eqs. (5.2.9) and (5.2.12) imply that $\theta(t) \xrightarrow[t \to +\infty]{} 0$ exponentially. Hence, we can define

$$\overline{\theta} = \lim_{t \to +\infty} \theta(t) = \lim_{t \to +\infty} \big(\theta(0) + \int_0^t \dot{\theta}(\tau) d\tau\big) = \theta(0) + \int_0^{+\infty} \dot{\theta}(\tau) d\tau \tag{5.2.14}$$

because the integral converges, see Eq. (5.2.13). Also,

$$|\theta(t) - \overline{\theta}| = |\int_t^{+\infty} \dot{\theta}(\tau) d\tau| \leq 2\, \frac{e^{-(\lambda_1 - \alpha)\, t}}{\lambda_1 - \alpha}\, c(\Omega + |\boldsymbol{\omega}|) \tag{5.2.15}$$

by Eqs. (5.2.9) and (5.2.12).

Possibly by rotating the fixed frame, suppose that $\overline{\theta} \neq 0, \pi$. Then Eq. (5.2.10) implies that $\dot{\varphi}$ tends to zero exponentially since, as above, it is

$$\overline{\varphi} = \varphi(0) + \int_0^{+\infty} \dot{\varphi}(\tau)\, d\tau, \tag{5.2.16}$$

$$|\varphi(t) - \overline{\varphi}| \leq \frac{2c(\Omega + |\boldsymbol{\omega}|)}{\inf_{\tau \geq t} |\sin\theta(\tau)|} \frac{e^{-(\lambda_1 - \alpha)\, t}}{\lambda_1 - \alpha} \tag{5.2.17}$$

which show the second of Eqs. (5.2.13).

Similarly Eqs. (5.2.12) and (5.2.11) imply that ω_3 approaches $\widehat{\omega}_3$ exponentially, as $t \to +\infty$. Hence, setting

$$\overline{\psi} = \psi(0) + \int_0^{+\infty} (\dot{\psi}(\tau) - \widehat{\omega}_3)\, d\tau, \tag{5.2.18}$$

one finds, by Eqs. (5.2.11) and (5.2.12), for $t \geq t_1$

$$\begin{aligned} &|\psi(t) - \overline{\psi} - \widehat{\omega}_3 t| = |\psi(0) + \int_0^{+\infty} \dot{\psi}(\tau) d\tau - \overline{\psi} - \widehat{\omega}_3 t| \\ &= |\int_t^{+\infty} (\dot{\psi}(\tau) - \widehat{\omega}_3)\, d\tau \leq \frac{2\, c(\Omega + |\boldsymbol{\omega}|)}{\inf_{\tau \geq t_1} |\sin\theta(\tau)|} \frac{e^{-(\lambda_1 - \alpha)\, t}}{\lambda_1 - \alpha} \end{aligned} \tag{5.2.19}$$

proving Eq. (5.2.13). Naturally, the time t_1 has to be chosen so that $\inf_{\tau \geq t_1} |\sin\theta(\tau)| \geq \frac{1}{2} |\sin\overline{\theta}| > 0$, say.

To prove Eq. (5.2.12) remark that from Eq. (5.1.19), multiplying the first equation by ω_1, and the second by ω_2 and adding the results, one finds

$$\frac{d}{dt}\frac{1}{2}(\omega_1^2 + \omega_2^2) \leq -(\lambda - \alpha)(\omega_1^2 + \omega_2^2), \qquad \text{hence} \tag{5.2.20}$$

$$\omega_1(t)^2 + \omega_2(t)^2) \leq \big(\omega_1(0)^2 + \omega_2(0)^2\big) e^{-2(\lambda_1 - \alpha)\, t} \tag{5.2.21}$$

Furthermore, setting $z \stackrel{def}{=} \omega_3 - \widehat{\omega}_3$, the third of the Eqs. (5.1.19) becomes

$$\begin{aligned} \dot{z} =& \dot{\omega}_3 - \lambda_1 z - \lambda_2'''\big(\omega_3^2 - \widehat{\omega}_3^2\big) - \big(\lambda_2' \omega_1^2 + \lambda_2''\, \omega_2^2\big)\omega_3 \\ =& \big(-\lambda_1 - \lambda_2' \omega_1^2 - \lambda_2'' \omega_2^2\big) z - \lambda_2'''\big(\omega_3^2 + \widehat{\omega}_3 \omega_3 + \omega_3^2\big) z \\ & - \widehat{\omega}_3 (\lambda_1' \omega_1^2 + \lambda_2'' \omega_2^2). \end{aligned} \tag{5.2.22}$$

Since the general solution to the equation

$$\dot{y} = f(t) u + g(t), \qquad t \geq 0, \tag{5.2.23}$$

is, $\forall\, f, g \in C^\infty(\mathcal{R})$,

$$y(t) = y(0)\, e^{\int_0^t f(\tau) d\tau} + \int_0^t g(\tau) e^{\int_\tau^t f(\theta) d\theta}\, d\tau, \tag{5.2.24}$$

Eq (5.2.21) implies

$$\begin{aligned} z(t) = & z(0)\, e^{-\int_0^t \left(-\lambda_1 - \lambda_2'\omega_1^2 - \lambda_2''\omega_2^2) - \lambda_2'''(\omega_3^2 + \widehat{\omega}_3\omega_3 + \omega_3^2\right) d\tau} \\ & - \int_0^t \widehat{\omega}_3\,(\lambda_2'\omega_1^2 + \lambda_1''\omega_2^2) e^{-\int_\tau^t \left(\lambda_1 + \lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2'''(\omega_3^2 + \widehat{\omega}_3\omega_3 + \omega_3^2)\right) d\theta} d\tau. \end{aligned} \tag{5.2.25}$$

The functions which multiply $\lambda_2', \lambda_2'', \lambda_2'''$ are nonnegative therefore

$$\begin{aligned} |z(t) \le & |z(0)|e^{-\lambda_1 t} + \overline{\lambda}_2 |\widehat{\omega}_3|(\omega_1(0)^2 + \omega_2(0)^2) \int_0^t e^{-2(\lambda_1 - \alpha)\tau} e^{-\lambda_1(t-\tau)} d\tau \\ \le & e^{-(\lambda_1 - \alpha)t} \big(|z(0)| + \overline{\lambda}_2 \frac{|\widehat{\omega}_3|(\omega_1(0)^2 + \omega_2(0)^2)}{\lambda_1 - \alpha}\big) \\ \le & \big(|\widehat{\omega}_3| + |\omega_3(0)| + \frac{\overline{\lambda}_2}{\lambda_1 - \alpha} |\widehat{\omega}_3| \boldsymbol{\omega}(0)^2\big), \end{aligned} \tag{5.2.26}$$

by Eq. (5.2.21) if $\overline{\lambda}_2 \stackrel{def}{=} \max(\lambda_1', \lambda_2'')$.

Hence, Eq. (5.2.12) follows from Eqs. (5.2.26) and (5.2.21) with

$$c(x)^2 = \big(|\widehat{\omega}_3| + x + \frac{\overline{\lambda}_2}{\lambda_1 - \alpha} |\widehat{\omega}_3 x^2\big) \tag{5.2.27}$$

mbe

The analysis for $\alpha > \lambda_1$, is much more interesting and involves quite a few general ideas which will be discussed in the upcoming sections. The character of motion will change: for $\alpha \gg \lambda_1$ it will be described, asymptotically for $t \to +\infty$, by a behavior very different from the one seen so far, where the gyroscope sets itself in a state of uniform rotation around the axis $\mathbf{i}_3$, fixed in space.

5.2.1 Exercises

1. Suppose that $R = 0$ in Eq. (5.1.18). Show that for $\alpha < \lambda_1$, something analogous to the statement of Proposition 4 holds.

2. Same as Problem 1, for Eq. (5.1.19).

3. Show that for $\alpha > \lambda_1$, Eqs. (5.1.18) and (5.1.19) with $R = 0$ admit infinitely many stationary solutions, and find them.

4. Consider a gyroscope like the one in Eq. (5.1.14), but assume that the friction is linear, that the two little jets arranged along the $\mathbf{i}_2$ axis in $-\varrho\mathbf{i}_2$ or $+\varrho\mathbf{i}_2$ produce a thrust in the direction $\mathbf{i}_3$ equal to $f_1\mathbf{i}_1$, while the two jets on $\pm\varrho'\mathbf{i}_3$ produce a constant thrust $R\,\mathbf{i}_1$. Show that the equations of motion become

$$\begin{aligned} \dot{\omega}_1 = & -\lambda\,\omega_1 + \omega_2\omega_3 - \sigma\,\omega_3, \\ \dot{\omega}_2 = & -\lambda\,\omega_2 - \omega_1\omega_3 + \alpha, \\ \dot{\omega}_3 = & -\lambda\,\omega_3 + \sigma\,\omega_1 \end{aligned}$$

for suitably chosen α, σ and after a change of variables $\omega_i \to \omega_i \frac{J-I}{I}$. Find the stationary solutions for the above equation.

5. Set $\omega_3 = x, \omega_2 = z, \omega_1 = y$ and suppose that the friction is different for the different components of the angular velocity; i.e., suppose that the friction moment is $(-\lambda_1\omega_1, -\lambda_2\omega_2, -l_3\omega_3)$. Study the same problem as in Problem 4 with $\lambda_1 = 1, \lambda_2 = b, \lambda_3 = \sigma$, fixing $b = \frac{8}{3}$, $\sigma = 10$ ("Lorenz model").

6. Find whether an analogue of Proposition 4 holds for the equations in Problems 4 and 5 for some values of α.

7. Find the stationary solutions for the equations

$$\begin{aligned}
\dot\gamma_1 &= -2\gamma_1 + 4\gamma_2\gamma_3 + 4\gamma_4\gamma_5,\\
\dot\gamma_2 &= -9\gamma_2 + 3\gamma_1\gamma_3,\\
\dot\gamma_3 &= -5\gamma_3 - 7\gamma_1\gamma_2 + \alpha,\\
\dot\gamma_4 &= -5\gamma_4 - \gamma_1\gamma_5,\\
\dot\gamma_5 &= -\gamma_5 - 3\gamma_1\gamma_4.
\end{aligned}$$

Using the same method of the proof of Proposition 4, for α small, find a proof of the statement analogous to that appearing in Proposition 4, Eq. (5.2.12) ("five-mode approximation to the Navier-Stokes equations on $\mathcal{T}^2$").

8. Same as Problem 7 for the equations

$$\begin{aligned}
\dot\gamma_1 &= -2\gamma_1 + 4\sqrt{5}\gamma_2\gamma_3 + 4\sqrt{5}\gamma_4\gamma_5,\\
\dot\gamma_2 &= -9\gamma_2 + 3\sqrt{5}\gamma_1\gamma_3,\\
\dot\gamma_3 &= -5\gamma_3 - 7\sqrt{5}\gamma_1\gamma_2 - 9\gamma_1\gamma_7 + \alpha,\\
\dot\gamma_4 &= -5\gamma_4 - \sqrt{5}\gamma_1\gamma_5,\\
\dot\gamma_5 &= -\gamma_5 - 3\sqrt{5}\gamma_1\gamma_4 - 5\gamma_1\gamma_6\\
\dot\gamma_6 &= -\gamma_6 - 5\gamma_1\gamma_5,\\
\dot\gamma_7 &= -5\gamma_7 - 9\gamma_1\gamma_3,
\end{aligned}$$

("seven-mode truncation of the Navier-Stokes equations on $\mathcal{T}^2$").

5.3 Attractors and Stability

For $\alpha > \lambda_1$, the motions of the model considered in §5.2 will exhibit a behavior qualitatively different from that seen for $\alpha < \lambda_1$. It is therefore convenient to introduce some notions well suited to discuss various results in suggestive and agile language.

The notions on stability and attractors that will be introduced can be subjected to the same critiques already presented in Chapter 2 when we introduced similar notions; i.e., they should not be taken too seriously as absolute definitions. Usually everyone, motivated by their own scopes, ideas, and needs, introduce their own definitions and it makes no sense to insist on a standard nomenclature, as much as it makes no sense to agree once and for all on the choice of the units of measure of the various physical entities. Here we shall

choose some significant definitions and not discuss alternative definitions, recalling that in applications the "correct" notions of stability and attractivity will be determined by the applications themselves.

In this and in the following sections, autonomous differential equations in $\mathcal{R}^d$ of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{5.3.1}$$

will be considered, supposing that the solutions have bounded trajectories, see Definition 3, §2.5, p.28, i.e., that the solution flow S_t to Eq. (5.3.1) has the property that there exists a function $\mu : \mathcal{R}_+ \to \mathcal{R}_+$ such that

$$|S_t(\mathbf{u})| \le \mu(|\mathbf{u}|), \qquad \forall\, t \ge 0, \forall\, \mathbf{u} \in \mathcal{R}^d. \tag{5.3.2}$$

Proposition 1, §5.2, p.369, shows that Eq. (5.1.19) has this property with $\mu(|u|) = |u| + \Omega$.

The first interesting notion is that of a stable set.

1 Definition. *Consider the flow S_t solving, for $t > 0$, a differential equation in $\mathcal{R}^d$ like Eq. (5.3.1), with bounded trajectories. If $A \subset \mathcal{R}^d$, we denote $S_t(A)$ the set of the points $\mathbf{u}$ having the form $\mathbf{u} = S_t(\mathbf{w})$ for some $\mathbf{w} \in A$.*
A set A will be called "invariant" for Eq. (5.3.1) [or for the motions of Eq. (5.3.1) or for its trajectories] if

$$S_t(A) \subset A, \qquad \forall\, t \ge 0, \tag{5.3.3}$$

i.e., A is invariant if the trajectories originating in A develop, entirely, within A. If the inclusion in (5.3.3) holds also for $t < 0$, the set A will be called "bi-invariant".
An invariant, set A will be called "stable" for the evolution described by Eq. (5.3.1) if every neighborhood U of A contains a neighborhood V such that

$$S_t(V) \subset U, \qquad \forall\, t \ge 0, \tag{5.3.4}$$

i.e., A is stable if motions starting sufficiently close to A do not go too far from it.

Examples
(1) The equation of the harmonic oscillator,

$$\dot{x} = -y, \quad \dot{y} = x, \tag{5.3.5}$$

is an equation in $\mathcal{R}^2$ such that every circle around the origin is invariant and stable.
(2) Proposition 1, §5.2, relative to the gyroscope equation (5.1.19), provides another example. Equation (5.2.1) says that the ball with radius 2Ω is invariant. From Eq. (5.2.1), it also follows that it is stable.

Another notion, closely related to the above, is that of attractor.

2 Definition. *A closed set $A \subset \mathcal{R}^d$, invariant for the evolution associated to Eq. (5.3.1), is called an "attractor" for the motions of Eq. (5.3.1) if there exists an open set $U \supset A$ such that*

$$\lim_{t\to+\infty} d(S_t(\mathbf{u}), A) = 0, \quad \forall \mathbf{u} \in U, \tag{5.3.6}$$

where $d(x, A) =$ (distance of x from A) and the set U is said to be a "partial basin of attraction" for A.
The union of all the partial basins of attraction will be called the "attraction basin" of A and denoted as $B(A)$.
An attractor A is called minimal if it does not contain any proper subset which is also an attractor.
A partial basin of attraction U for an attractor A will be called "normal" if for every $\mathbf{u} \in U$ there is at least one point $\pi(\mathbf{u}) \in A$ such that

$$\lim_{t\to+\infty} d(S_t(\mathbf{u}), S_t(\pi(\mathbf{u}))) = 0, \tag{5.3.7}$$

and the point $\pi(\mathbf{u})$ will be called a "projection" of $\mathbf{u}$ on A.

Examples and Observations
(1) The ball with radius 2Ω, as well as that with radius Ω, are attractors for Eq. (5.1.19). The first statement follows from Eq. (5.2.2), while the second can be deduced from the remark following Eq. (5.2.5) by slightly improving it (exercise).
(2) For $\alpha < \lambda$ the point $\widehat{\boldsymbol{\omega}}$ is an attractor for Eq. (5.1.19) as is shown by Eq. (5.2.12). Its basin is all of $\mathcal{R}^3$, and it is a normal basin. Clearly, every basin of attraction for an attractor consisting of just one point is normal for it.
(3) The unit circle is an attractor for the solutions of the equation in $\mathcal{R}^2$:

$$\dot{x} = -\frac{1}{2}x\,(x^2 + y^2 - 1), \qquad \dot{y} = -\frac{1}{2}y\,(x^2 + y^2 - 1), \tag{5.3.8}$$

In fact, by multiplying the first of Eqs. (5.3.8) by x the second by y and adding the results,

$$\frac{d}{dt}\frac{x^2 + y^2}{2} = -(x^1 + y^2 - 1)\frac{x^2 + y^2}{2} \tag{5.3.9}$$

Setting $\varrho = x^2 + y^2$, this becomes $\dot{\varrho} = -\varrho(\varrho - 1)$, implying, if $\varrho(0) \neq 0$,

$$\frac{\varrho(t) - 1}{\varrho(t)} = \frac{\varrho(0) - 1}{\varrho(0)}\,e^{-t} \tag{5.3.10}$$

hence $\lim_{t\to+\infty} \varrho(t) = 1$ and the attraction basin for the unit circle consists of $\mathcal{R}^2/\{\mathbf{0}\}$. The basin is normal because the point $(x, y) \neq \mathbf{0}$ has projection

$$\pi(x, y) = \Big(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}}\Big) \tag{5.3.11}$$

on it and, in this case, the projection on the attractor is unique.

As an exercise one can look at the trajectories of Eqs. (5.3.8) and at the geometrical meaning of Eq. (5.3.11). The unit circle is an attractor consisting of fixed points; it is also minimal.

(4) In general, it is not true that an attractor is a stable set.

To obtain some understanding of the mechanism (somewhat pathological, in fact) by which a point may be attractive without being stable, consider the unit circle S^1 in $\mathcal{R}^2$ and let $f \in C^\infty(S^1)$ be a function described as $\theta \to f(\theta)$, where $\theta \in [0, 2\pi]$ parameterizes a point on S^1. Suppose that $f(\theta) > 0$, $\forall\, \theta \in (0, 2\pi)$ and $f(0) = 0 = f(2\pi)$; then, by the Taylor expansion, one realizes that $1/f(\theta)$ is not summable to either the right or to the left of 0. Consider the equation

$$\dot{\theta} = f(\theta) \tag{5.3.12}$$

as an equation of motion of a point moving on S^1, interpreting the angle θ, in Fig. 5.1, as the position.

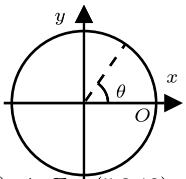


Figure 5.1: Illustration of remark (4) via Eq. (5.3.12).

It appears immediately that since $f(0) = 0$, the point $\theta = 0$ is an equilibrium position for Eq. (5.3.12). But if $\theta_0 > 0$, then $S_t(\theta_0) = \theta(t)$ increases with t, because $f > 0$ and f vanishes only for $\theta = 0$ or $\theta = 2\pi$, and it takes an infinite amount of time to reach 2π. This is so because the time needed to reach 2π starting from $\theta_0 < 2\pi$ is $\int_{\eta_0}^{2\pi} \frac{d\theta}{f(\theta)} = +\infty$ since $f(\theta)^{-1}$ is not integrable. However, in a finite time, $\theta(t)$ reaches any other position $\theta' \in (\theta_0, 2\pi)$ (as $\int_{\theta_0}^{\theta'} \frac{d\theta}{f(\theta)} < +\infty$). Hence,

$$\lim_{t \to +\infty} \theta(t) = 2\pi, \quad \forall\, \theta_0 \in (0.2\pi). \tag{5.3.13}$$

All circle points evolve counterclockwise towards 2π, reaching it from the left, with the obvious exception of the points $\theta_0 = 0$ and $\theta_0 = 2\pi$. Next, let $\mathbf{f}$ be an $\mathcal{R}^2$-valued function in $C^\infty(\mathcal{R}^2)$ which in a circular annulus U around the unit circle has the value

$$\mathbf{f}(x, y) = \big(- y f(\theta) - \frac{x}{2}\,(x^2 + y^2 - 1), x f(\theta) - \frac{y}{2}\,(x^2 + y^2 - 1)\big) \tag{5.3.14}$$

if (r, θ) are the polar coordinates of (x, y).

The equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ associated with Eq. (5.3.14) can be written in polar coordinates and for $(x, y) \in U$:

$$\dot{\theta} = f(\theta), \qquad \frac{d}{dt} r^2 = -r^2(r^2 - 1) \tag{5.3.15}$$

and the second relation shows that the set U is invariant and that the unit circle is an attractor. The first of Eqs. (5.3.15) shows that the point $\theta = 0, r = 1$ is a minimal attractor on the unit circle, which is unstable since arbitrarily close to it, there are points reaching it after going as far as ~ 2 away (e.g., the point $r = 1, \theta = \varepsilon > 0$), i.e., after traveling a distance approximately equal to the circle diameter.
(5) As the reader may guess, the problem of finding the basin of attraction of an attractor is a difficult problem. Very often it is only possible to determine some partial basins of attraction. The same remark applies to the determination of the minimal attractors.
In many applications, knowing partial domains of attraction or non minimal attractors is sufficient and the knowledge of such "global properties" as the maximal basins or the minimal attractors are not needed.
(6) It is convenient not to require that a partial basin of attraction U for A be invariant. This may rightly be considered a natural requirement; note, however, that $V = U \cup_{t \geq 0} S_t(U)$ is an open invariant basin of attraction for A, i.e., any partial basin of attraction for A is contained inside an invariant partial basin of attraction. The total basin $B(A)$ is obviously invariant.
(7) If the differential equation (5.3.1) is also normal in the past, see p.28, it is possible to construct $B(A)$ from a partial basin U for A as $B(A) = \cup_{t \in \mathcal{R}} S_t(U)$.

The question of the normality of a basin U for an attractor A is obviously quite important. For simplicity, assume A bi-invariant.

Intuitively, the normality of U with respect to A depends on two factors: the speed of approach of the points of U to A and the speed of reciprocal separation of two points in A. One can expect that U is normal with respect to A if the speed of reciprocal separation of two points in A is much smaller than the speed of approach to A by the points of U.

To make precise this intuitive idea, let us introduce some new concepts.

3 Definition. *Let U be a partial attraction basin for an attractor A for Eq. (5.3.1). We define the "attraction modulus" of A for U (or the "attractor strength") as the function*

$$d_U(t) = \sup_{\substack{\mathbf{u} \in U \\ \tau \geq t}} d(S_\tau(\mathbf{u}), A), \tag{5.3.16}$$

and $d_U(t)$ may be $+\infty$. Note that $d_U(t)$ decreases monotonically with t.

Together with this notion, it is convenient to introduce another notion measuring how quickly two points on A can separate from each other. Note that from the regularity theorem for differential equations, see §2.4, it follows that if Γ is a bounded closed set, the quantity

$$\sup_{\substack{\mathbf{x},\mathbf{x}'\in\Gamma \\ \mathbf{x}\neq\mathbf{x}'}} \frac{|S_t(\mathbf{x}) - S_t(\mathbf{x}')|}{|\mathbf{x}-\mathbf{x}'|} = m_t(\Gamma) \tag{5.3.17}$$

is finite for all $t > 0$ and bounded on every finite interval $[0, T]$, $T > 0$. It can be naturally called the "maximal expansion rate" for Eq. (5.3.1) relative to $t \in \mathcal{R}$ and to $\Gamma \subset \mathcal{R}^d$. To this notion, the following definition is related.

4 Definition. *Let A be a bi-invariant attractor for Eq. (5.3.1) which is not a single point. The "uniform coefficient of maximal expansion" for Eq. (5.3.1) on A will be defined as the quantity*

$$M_t(A) = \sup_{\substack{\mathbf{x}\neq\mathbf{x}'\in A \\ |\tau|\leq t}} \frac{|S_\tau(\mathbf{x}) - S_\tau(\mathbf{x}')|}{|\mathbf{x}-\mathbf{x}'|} = \sup_{\tau\leq t} m_\tau(A), \tag{5.3.18}$$

Note that $M_t(A)$ is monotonically increasing with t for $t > 0$.

Observations.
(1) The normality and boundedness assumptions on trajectories of Eq. (5.3.1), made at the beginning of this section, do not guarantee existence of global solutions in the past for all initial data.[5] Hence, it is important to stress that in Eq. (5.3.18) A is bi-invariant and negative times are also involved.
(2) Even if A is bounded, so that $|S_\tau(x) - S_\tau(x')| \leq \{\text{diameter of } A\}$ for all τ, the function $M_t(A)$ can increase very rapidly with t. A simple though rather trivial example is the following. Let $f \in C^\infty(\mathcal{R})$ be such that

$$\begin{aligned} f(x) =& x && \text{if } |x| < \frac{1}{2} \\ f(x) =& -x(x^2 - 1) && \text{if } |x| < 1. \end{aligned} \tag{5.3.19}$$

Then the interval $[-1, 1]$ is an attractor for the solutions of the differential equation $z = f(x)$ and

$$M_t([-1, 1]) \geq e^t. \tag{5.3.20}$$

Eq. (5.3.20) follows by considering the evolutions of $x_0 = 0$ and $x_1 = \varepsilon \neq 0$.
(3) By definition, $M_t(A) \geq 1$. When A is a single point, we shall set $M_t(A) = 1$.
(4) If A is a periodic orbit with minimal period $T > 0$, then $M_t(A)$ is bounded in t and $M_t(A) \leq M_T(A), \forall\, t \geq 0$.

The following proposition makes quantitative the idea discussed above about the normality of an attraction basin U for an attractor A of Eq. (5.3.1). It provides a sufficient, though by no means necessary, condition for the normality of a basin.

5 Proposition. *Let A be a bounded bi-invariant attractor for Eq. (5.3.1) and let U be an attraction basin for A. Assume the existence of $C > 0, \varepsilon > 0$ such that for all $t \geq 0$:*

[5] For instance, the differential equation $\dot{x} = -\frac{1}{2}x^3$ is normal in the future but not in the past; its solutions cannot be extended beyond $t_0 = -x(0)^{-2}$.

$$M_{t+1}(A)^2 d_U(t) < \frac{C}{(1+t)^{1+\varepsilon}}. \tag{5.3.21}$$

Then U is normal for A.
If A is a periodic trajectory, it is normal if there is a $C_>0$ such that

$$d_U(t) < \frac{C_1}{(1+t)^{1+\varepsilon}}. \tag{5.3.22}$$

Observations.
(1) Note that the statement concerning the periodic orbits is a consequence of the general statement. In fact, if A is a periodic orbit, it is clear that $M_t(A)$ is bounded, see Observation (4), to Definition 4 above.
(2) Equation (5.3.21) implies the existence of a constant C_2 such that

$$m_\tau(A) \le C_2, \qquad \forall\, \tau \in [-1,1]. \tag{5.3.23}$$

It also implies

$$\begin{aligned} &\text{diameter of } U < (2d_U(0) + \text{diameter of } A)\\ &\le (2C + \text{diameter of } A) < +\infty \end{aligned} \tag{5.3.24}$$

PROOF. Let $t_n = n, n = 0, 1, ...,$ and let $\mathbf{x} \in U$. Let $\mathbf{a}_n \in A$ be a point with minimal distance from $S_n(\mathbf{x})$, among the points of A. The natural idea is that a projection $\mathbf{a}(\mathbf{x})$ of $\mathbf{x}$ can be defined as

$$\pi(\mathbf{x}) = \lim_{n\to+\infty} S_{-n}(\mathbf{a}_n). \tag{5.3.25}$$

To prove the existence of the above limit, let us compare $S_{-n}(\mathbf{a}_n)$ with $S_{-n-1}(\mathbf{a}_{n+1})$, assuming that A is not a single point (a case in which everything becomes trivial). Let U be the closure of U, bounded by Eq. (5.3.24). By the remark after Eq. (5.3.17), $\sup_{\tau\in[0,1]} m_\tau(\overline{U}) \le \mu < +\infty$. By Eq. (5.3.21),

$$\begin{aligned} &|S_{-n}(\mathbf{a}_n) - S_{-(n+1)}(\mathbf{a}_{n+1})| \le M_{n+1}(A)\,|S_1(\mathbf{a}_n) - \mathbf{a}_{n+1}|\\ &\le M_{n+1}(A)\,(|S_1(\mathbf{a}_n) - S_{n+1}(\mathbf{x})| + |S_{n+1}(\mathbf{x}) - \mathbf{a}_{n+1}|)\\ &\le M_{n+1}(A)\,(|S_1(\mathbf{a}_n) - S_1 S_n(\mathbf{x})| + d_U(n+1))\\ &\le M_{n+1}(A)\,(m_1(U) d_U(n) + d_U(n+1))\\ &\le M_{n+1}(A) d_U(n)(1 + m_1(U)) \le \frac{C\,(1+m_1(U))}{(1+n)^{1+\varepsilon}} M_{n+1}(A)^{-1} \end{aligned} \tag{5.3.26}$$

Hence, the series $\sum_{n=0}^{\infty} |S_{-n}(\mathbf{a}_n) - S_{-n-1}(\mathbf{a}_{n+1})|$ converges and, therefore, the limit of Eq. (5.3.25) exists. It also verifies

$$
\begin{aligned}
|\pi(\mathbf{x}) - S_{-n}(\mathbf{a}_n)| &= |\pi(\mathbf{x}) - \mathbf{a}_0 - \sum_{k=1}^{n} \big(S_{-k}(\mathbf{a}_k) - S_{-(k-1)}(\mathbf{a}_{k-1})\big)| \\
&\le C\, M_{n+1}(A)^{-1}(1+m_1(U)) \sum_{h=n+1}^{\infty} \frac{1}{h^{1+\varepsilon}} \xrightarrow[n\to\infty]{} 0.
\end{aligned} \tag{5.3.27}
$$

We now compare $S_n(\pi(\mathbf{x}))$ with $S_n(\mathbf{x})$:

$$
\begin{aligned}
&|S_n(\pi(\mathbf{x})) - S_n(\mathbf{x})| \le |S_n(\pi(\mathbf{x})) - \mathbf{a}_n| + |\mathbf{a}_n - S_n(\mathbf{x})| \\
&\le |S_n(\pi(\mathbf{x})) - S_n(S_{-n}(\mathbf{a}_n))| + d_U(n) \le M_n(A)|\pi(\mathbf{x}) - S_{-n}(\mathbf{a}_n)| + d_U(n) \\
&\le C\,(1+m_1(U)) \sum_{h=n+1}^{\infty} \frac{1}{h^{1+\varepsilon}} + d_U(n) \xrightarrow[n\to\infty]{} 0.
\end{aligned} \tag{5.3.28}
$$

Finally, if $t = n+\tau, \tau \in (0,1)$, is large enough,

$$
\begin{aligned}
&|S_t(\mathbf{x}) - S_t(\pi(\mathbf{x}))| = |S_\tau S_n(\mathbf{x}) - S_\tau S_n(\pi(\mathbf{x}))| \\
&\le m_\tau(U)|S_n(\mathbf{x}) - S_n(\pi(\mathbf{x}))| \le \mu\,|S_n(\mathbf{x}) - S_n(\pi(\mathbf{x}))|
\end{aligned} \tag{5.3.29}
$$

because $S_n(\mathbf{x}) \in U$ for n large enough. Since the right-hand side of Eq. (5.3.29) approaches zero, by Eq. (5.3.28) the proposition is proved. mbe

5.3.1 Exercises

1. Investigate the normality of some basins of partial attraction for the attractors associated with the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ in $\mathcal{R}^2$:

$$
\mathbf{f}(x,y) = \Big(-y\,(x^2+y^2-1) - \frac{x}{2}\psi(x^2+y^2), x\,(x^2+y^2-1) - \frac{y}{2}\psi(x^2+y^2)\Big),
$$

where $\psi \ge 0$ is a C^∞ function of its argument vanishing in 1 only. Show that the normality of the attractor is related to the convergence of the integral $\int \frac{dr^2}{r^2}\,\frac{r^2-1}{\psi(r^2)}$ near $r=1$.

2. An attractor may be minimal and non connected. Find an example. (Hint: Starting from Observation (4) to Definition 2, 377, improve the idea, i.e., take $f(\theta)$ vanishing not only in 0 and 2π, but also in π, positive elsewhere, so that the integral $\int f(\theta)\frac{1}{d\theta}$ diverges near 0 and π.)

3. Consider a Hamiltonian system with ℓ degrees of freedom, integrable on some region W of its phase space. Show that each of the tori covering W is an invariant set for the Hamiltonian flow. Each is stable, but none are attractive.

4. In the context of Problem 3, note that the invariant tori in W having pulsations $\boldsymbol{\omega}$ with rational components are covered by periodic orbits. Show that none of these orbits are stable if the Jacobian matrix $J_{ij} = (\frac{\partial \omega_i}{\partial A_j}(\mathbf{A})$ has a non vanishing determinant in W. (*Hint.* As close as we wish to a given "rational" torus, there must be one with rationally independent components if $\det J \neq 0$ (use the implicit function theorem, or see Problem 15, §5.10, p.478). Every point on such "irrational torus" evolves covering it densely, and this implies instability because .... On the contrary if $\frac{\partial \boldsymbol{\omega}}{\partial \mathbf{A}} = \mathbf{0}$, the periodic orbits, when existing, are stable.)

5. Let A_1, A_2 be two attractors with partial basins U_1, U_2, respectively. Show that $A_1 \cap A_2$ is an attractor with partial basin $U_1 \cap U_2$, if $A_1 \cap A_2 \neq \emptyset$.

6. Show that if the set of the attractors contained in a given bounded attractor A is finite then there is a minimal attractor in A.

7. Find an example of "an attractor without minimal attractors". (*Hint:* Let $f \in C^\infty(\mathcal{R})$ be everywhere positive for $x < 0$ except at the points $x_j = -\frac{1}{j}$, $j = 1, 2, \ldots$, where it vanishes (so that $\int \frac{dx}{f(x)}$ does not converge near any of the x_j). Suppose, also, that $f(x) < 0$ for $x > 0$. Then $\dot{x} = f(x)$ admits $[x_j, 0]$ as attractors; however, it has no minimal attractors because $\{0\}$ is not an attractor.)

8. Show that the bi-invariance assumption is essential in proposition 5. (*Hint:* Consider $\dot{x} = -x$, $A = [-1, 1]$ and show that A is not normal.)

9. Show that every bounded attractor A contains a bi-invariant attractor $\widetilde{A}$. (*Hint:*$\widetilde{A} = \cap_{t \geq 0} S_t(A)$.)

5.4 The Stability Criterion of Lyapunov

Consider a differential equation, like Eq. (5.3.1), with bounded trajectories. A simple and useful criterion for the stability of one of its stationary solutions ("fixed points") is the following proposition ("Lyapunov's theorem").

6 Proposition. *Let* $\mathbf{x}_0$ *be an equilibrium point for Eq. (5.3.1),* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$*, with* $\mathbf{f} \in C^\infty(\mathcal{R}^d)$*:*

$$\mathbf{f}(\mathbf{x}) = (f^{(1)}(\mathbf{x}), \ldots, f^{(d)}(\mathbf{x})) \tag{5.4.1}$$

and define the "stability matrix" (or "Lyapunov matrix")

$$L_{ij} = \frac{\partial f^{(i)}(\mathbf{x}_0)}{\partial x_j}, \qquad i, j = 1, \ldots, d. \tag{5.4.2}$$

If the eigenvalues of L*, i.e., the solutions of the* d*-th degree equation in* λ

$$\det(L - \lambda) = 0 \tag{5.4.3}$$

(see Appendix E), have a negative real part, then $\mathbf{x}_0$ *is stable and is locally attractive with exponential strength.*[6]

If at least one of the eigenvalues has a positive real part, then $\mathbf{x}_0$ *is unstable.*

Observations.
(1) More precisely, if all the eigenvalues $\lambda_1, \ldots, , \lambda_d$ of L have a negative real part, there exists $t_0 > 0$ ("halving time") and $\varrho > 0$ such that for $|\mathbf{x}_0 - \mathbf{w}| < \varrho$, one has

$$d(S_t(\mathbf{w}), \mathbf{x}) \leq 2 \cdot 2^{-\frac{t}{t_0}} |\mathbf{w}|, \qquad \forall\, t \geq t_0. \tag{5.4.4}$$

(2) The reason why the above proposition is true and natural is made clear by the analysis of the "linear case", i.e., by the analysis of Eq. (5.3.1) with

[6] i.e. there is a small enough neighborhood U of $\mathbf{x}_0$ which is a partial basin of attraction for $\mathbf{x}_0$ with an exponential strength of attraction, see Definition 3, p.378.

$$f^{(i)}(\mathbf{x}) = \sum_{j=1}^{d} L_{ij} x_j = (L\mathbf{x})_j. \tag{5.4.5}$$

In this case, $\mathbf{x}_0 = 0$ is a stationary point for the equation; the equation itself can now be written as

$$\dot{\mathbf{x}} = L\mathbf{x}, \tag{5.4.6}$$

and its stability matrix is just L. As seen in the problems of §2.2-§2.6, one can look for d linearly independent solutions of Eq. (5.4.6) having the form

$$\mathbf{x}(t) = e^{\lambda\, t}\, \mathbf{v} \tag{5.4.7}$$

Such a solution exists if there exists $\mathbf{v} \neq \mathbf{0}$ such that

$$L\mathbf{v} = \lambda \mathbf{v} \tag{5.4.8}$$

If we assume that the d-th degree algebraic equation for λ, $\det(L - \lambda) = 0$, has d pairwise distinct roots $\lambda_1, \ldots, \lambda_d$ and if $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$ are the associated eigenvectors of Eq. (5.4.8), it is well known that $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$ are linearly independent (see Appendix E, p.523.) Then the function of $t \in R$:

$$\mathbf{x}(t) = \sum_{j=1}^{d} \alpha_j\, e^{\lambda_j\, t}\, \mathbf{v}^{(j)} \tag{5.4.9}$$

is, for every choice of $\alpha_1, \ldots, \alpha_d \in \mathcal{C}$, a solution to Eq. (5.4.6).

By the linear independence of the vectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$, by suitably fixing the coefficients $\alpha_1, \ldots, \alpha_d$ one can impose that Eq. (5.4.9) verifies any preassigned initial condition. Hence, Eq. (5.4.9) is the most general solution of Eq. (5.4.6). If $\mathcal{R}\mathrm{e}\, \lambda_i < 0$, $i = 1, \ldots, d$, it is clear that

$$|\mathbf{x}(t)| \leq e^{-\nu t} \sum_{j=1}^{d} |\alpha_j|\, |\mathbf{v}^{(j)}|, \qquad \forall t \geq 0 \tag{5.4.10}$$

where $-\nu = \max_{i=q,\ldots,d} \mathcal{R}\mathrm{e}\, \lambda_i < 0$; hence, the origin is an attractor with basin $\mathcal{R}^d$ itself. Every bounded sphere is attracted by the origin with exponential strength, by Eq. (5.4.10).

If instead $\mathcal{R}\mathrm{e}\, \lambda_1 > 0$ and $\mathcal{I}\mathrm{m}\, \lambda_1 \neq 0$, say, and if $\lambda_2 = \overline{\lambda_1}, \mathbf{v}^{(2)} = \overline{\mathbf{v}^{(1)}}$ (the bar denotes complex conjugation),[7] it is clear that by (5.4.9), the initial datum $\varepsilon\, (\mathbf{v}^{(1)} + \overline{\mathbf{v}^{(1)}})$ evolves into

$$2\, \varepsilon\, e^{\mathcal{R}\mathrm{e}\, \lambda_1\, t}\, \mathcal{R}\mathrm{e}\, (\mathbf{v}^{(1)} e^{i \mathcal{I}\mathrm{m}\, \lambda_1\, t}). \tag{5.4.11}$$

[7] Since L is a real matrix its eigenvalues appear in complex-conjugate pairs or are real. Similarly, the eigenvectors can be chosen to be either real or appearing in complex-conjugate pairs corresponding to complex-conjugate eigenvalues.

Hence, arbitrarily close to the origin, there are points evolving indefinitely far away from the origin. Therefore, O not only does not attract, but it is unstable.

The following proof will reduce the nonlinear case to the linear one. If $\mathcal{R}\text{e}\,\lambda_i < 0$, $i = 1, 2, \ldots, d$, one shows that if a point is close enough to the origin, then the nonlinear terms of $\mathbf{f}$ can initially be neglected for the purposes of studying the equation of the motion and, by the preceding argument, the point starts approaching O. Therefore, the nonlinear terms become even less important and, more and more precisely, the system will move as if it were subject to a linear equation.

If $\mathcal{R}\text{e}\,\lambda_1 > 0$, on the contrary, O cannot be stable because the initial datum $\varepsilon\,(\mathbf{v}^{(1)} + \overline{\mathbf{v}^{(1)}})$ moves away from the origin, if ε is small enough, at least as much as needed so that the nonlinear terms of the equation become sizeable. This suffices to exclude stability of the origin, even though it cannot exclude its attractivity (since the point could go far from O in the $\mathbf{v}^{(1)}, \overline{\mathbf{v}^{(1)}}$ plane (roughly)) and, then, under the influence of nonlinearity, it could come back towards 0 along a direction i where $\mathcal{R}\text{e}\,\lambda_i < 0$, except, of course, when $\mathcal{R}\text{e}\,\lambda_i > 0$, for all $i = 1, \ldots, d$. The reader will recognize the above ideas in the following proof.

PROOF. Let U_R be a radius R ball centered at the origin. Assuming that $\mathcal{R}\text{e}\,\lambda_i < 0$, $i = 1, \ldots, d$, we must determine ϱ_0 so that the evolution $t \to S_t(\mathbf{w})$ of an initial datum $\mathbf{w} \in U_{\varrho_0}$ develops, $\forall\, t \geq 0$ in U_R: $S_t(\mathbf{w}) \in U_R, \forall\, t \geq 0$.

For simplicity, suppose that $\lambda_1, \ldots, \lambda_d$ are pairwise distinct. The reader can think of the general case as a problem (basically, it is just an algebraic problem). Proceed as in the small oscillations theory of §2.14, Proposition 20, p.65, and write Eq. (5.3.1), assuming, without loss of generality, that $\mathbf{x}_0 = \mathbf{0}$:

$$\dot{\mathbf{x}} = L\mathbf{x} + (\mathbf{f}(\mathbf{x}) - L\mathbf{x}) \equiv L\mathbf{x} + \mathbf{N}(\mathbf{x}) \tag{5.4.12}$$

where $\mathbf{N}$ is an $\mathcal{R}^d$-valued $C^\infty(\mathcal{R}^d d)$ function with a second-order zero in O. By Taylor's theorem, see Appendix B, given $R > 0$, there is a constant C_R such that

$$|\mathbf{N}(\mathbf{x})| \leq C_R |\mathbf{x}|^2, \qquad \forall\, \mathbf{x} \in U_R \tag{5.4.13}$$

Consider Eq. (5.4.12) as an equation in which $\mathbf{N}(\mathbf{x}(t))$ is thought of as a known function of t, $\forall\, t \geq 0$. Then a particular "solution" would be

$$\mathbf{p}(t) = \int_0^t \sum_{i=1}^d e^{\lambda_i\,(t-\tau)} \alpha_i(\mathbf{N}(\mathbf{x})(\tau))\, \mathbf{v}^{(i)}\, d\tau \tag{5.4.14}$$

where, in general, given $\mathbf{w} \in \mathcal{R}^d$ we shall set

$$\mathbf{w} = \sum_{j=1}^d \alpha_j(\mathbf{w})\, \mathbf{v}^{(j)}. \tag{5.4.15}$$

Since $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$ is a basis in $\mathcal{C}^d$, such a representation is possible and defines the coefficients $\alpha_i(\mathbf{w})$ (which, in general, may be complex even for real $\mathbf{w}$); and, furthermore, there is a constant A such that

$$\sum_{j=1}^{d} |\alpha_j(\mathbf{w})| \leq A\,|\mathbf{w}|. \tag{5.4.16}$$

We shall suppose to have chosen the vectors $\mathbf{v}^{(i)}$ so that $|\mathbf{v}^{(i)}| \equiv 1$, $i = 1, \ldots, d$, which implies that $A \geq 1$. Then the solution to Eq. (5.4.12), $t \to \mathbf{x}(t)$, $t \geq 0$, with the initial datum $\mathbf{w}$ will be

$$S_t(\mathbf{w}) = \sum_{i=1}^{d} \alpha_i(\mathbf{w})\, e^{\lambda_i\, t}\, \mathbf{v}^{(i)} + \int_0^t \sum_{i=1}^{d} e^{\lambda_i\,(t-\tau)} \alpha_i(\mathbf{N}(S_\tau(\mathbf{w}))(\tau))\, \mathbf{v}^{(i)}\, d\tau. \tag{5.4.17}$$

The boundedness assumption on the trajectories implies existence of $\mu(R) < +\infty$ such that $|S_t(\mathbf{w})| \leq \mu(R), \forall\, t \geq 0, \forall\, \mathbf{w} \in U_R$. Then, setting, $\forall\, \varrho \leq R$,

$$D_\varrho(t) = \max_{\substack{0 \leq \tau \leq t \\ |\mathbf{w}| \leq \varrho}} |S_\tau(\mathbf{w})|, \tag{5.4.18}$$

one deduces from Eqs. (5.4.17), (5.4.18), (5.4.16), and (5.4.13):

$$|S_\tau(\mathbf{w})| \leq e^{-\nu t} A|\mathbf{w}| + A \int_0^t e^{-\nu\,(t-\tau)} C_{\mu(R)} D_\varrho(t)^2\, d\tau \leq A\varrho + \frac{AC_{\mu(R)}}{\nu} D_\varrho(t)^2, \tag{5.4.19}$$

where $\nu = \min_{i=1,\ldots,d} |\mathcal{R}\mathrm{e}\, \lambda_1|$. By the arbitrariness of t and by the monotonicity of $D_\varrho(t)$, as a function of t, Eq. (5.4.19) means that

$$D_\varrho(t) \leq A\varrho + \frac{AC_{\mu(R)}}{\nu} D_\varrho(t)^2, \tag{5.4.20}$$

i.e. if $4\frac{AC_{\mu(R)}}{\nu}\varrho < 1$, it must either be that

$$D_\varrho(t) \geq \frac{1 + \sqrt{1 - 4AC_{\mu(R)}\nu^{-1}\varrho}}{2AC_{\mu(R)}\nu^{-1}} \geq \frac{1}{2AC_{\mu(R)}\nu^{-1}} \tag{5.4.21}$$

or, if $K > 1$ is a suitably chosen constant (R-dependent).

$$D_\varrho(t) \leq \frac{1 - \sqrt{1 - 4AC_{\mu(R)}\nu^{-1}\varrho}}{2AC_{\mu(R)}\nu^{-1}} \leq K\,\varrho \tag{5.4.22}$$

If $|\mathbf{w}| \leq \varrho_0 < R$ and ϱ_0 is chosen so that

$$\varrho_0 = \frac{1}{2}\frac{1}{2AC_{\mu(R)}\nu^{-1}}, \tag{5.4.23}$$

we see that Eq. (5.4.22) must hold for all $t \geq 0$, by continuity, since for $t = 0$,

$$|\mathbf{w}| = D_\varrho(0) \leq \varrho_0. \tag{5.4.24}$$

Hence for all $\mathbf{w} \in T_{\varrho_0}$,

$$D_\varrho(t) \leq K\,|\mathbf{w}| \tag{5.4.25}$$

which implies that O is stable.

Attractivity of O is obtained via the autonomy of the Eq. (5.4.12) or (5.3.1). If in fact there is a time $t_0 > 0$ and a $\overline{\varrho} < \varrho_0$, [choose here ϱ_0 as given by Eq. (5.4.23) with $R = 1$, say], such that

$$|S_t(\mathbf{w})| \leq \frac{1}{2}|\mathbf{w}|, \qquad \forall\, t \geq t_0,\ \mathbf{w} \in U_{\overline{\varrho}}, \tag{5.4.26}$$

Then by the autonomy of the differential equation it is

$$|S_t(\mathbf{w})| \leq 2^{-n}|\mathbf{w}|, \qquad \forall\ t \geq nt_0,\, \mathbf{w} \in U_{\overline{\varrho}}, \tag{5.4.27}$$

as seen by iterating Eq. (5.4.26). Hence Eq. (5.4.27) implies

$$|S_t(\mathbf{w})| \leq 2 \cdot 2^{-\frac{t}{t_0}}|\mathbf{w}|, \qquad \forall\ t \geq t_0, \tag{5.4.28}$$

because $\frac{t}{t_0}$ is, in general, not an integer. It remains to check Eq. (5.4.26). The first of Eqs. (5.4.19), together with Eq. (5.4.25), implies

$$|S_t(\mathbf{w})|\lambda e e^{-\nu t}|\mathbf{w}| + \frac{AC_{\mu(1)}}{\nu}K^2|\mathbf{w}|^2 \leq |\mathbf{w}|(e^{-\nu t}A + \frac{AC_{\mu(1)}K^2\overline{\varrho}}{\nu}), \tag{5.4.29}$$

$\forall\,|\mathbf{w}| \leq \overline{\varrho}$ with $\overline{\varrho}$ arbitrary provided $\overline{\varrho} \leq \varrho_0$. If $\overline{\varrho}$ is chosen so small that $\frac{A}{\nu}C_{\mu(1)}K^2\overline{\varrho} < \frac{1}{4}$, it follows that

$$|S_t(\mathbf{w})| \leq (e^{-\nu t} + \frac{1}{4})|\mathbf{w}| \tag{5.4.30}$$

and Eq. (5.4.26) follows by choosing t_0 so that $Ae^{-\nu t_0} = \frac{1}{4}$, i.e. $t_0 = \frac{1}{\nu}\log 4A$.

The statement concerning the instability is left to the reader.

mbe

5.4.1 Exercises

1. Compute the Lyapunov matrix for the stationary points of the equation $\dot{x} = a\,x(1-x)$, $a \in \mathcal{R}$, and find for which values of a they are stable.

2. Consider the pendulum differential equation on $\mathcal{R}^2$: $\dot{x} = y$, $\dot{y} = -g\sin x$. Find the stationary points and compute their Lyapunov matrices, identifying the unstable ones. Find the explicit values of the eigenvalues of the Lyapunov matrix relative to all the stationary points and find the stable ones. (*Hint:* Stability cannot be decided on the basis of the Lyapunov's criterion; use energy conservation instead.)

3. Consider the Euler equations Eq. (4.11.32)-(4.11.34), p.312. Assume $I_1 < I_2 < I_3$ and compute the Lyapunov matrix of the stationary solutions different from $\boldsymbol{\omega} = \mathbf{0}$. Show that the only other stationary solutions are uniform rotations around either the $\mathbf{i}_1$ axis, or the $\mathbf{i}_2$ axis, or the $\mathbf{i}_3$ axis. Show that for solutions of this type the Lyapunov criterion does not exclude stability for the rotations around $\mathbf{i}_1$ and $\mathbf{i}_3$.

4. In the context of Problem 3, with $I_1 = I_2 = I$, $I_3 = J$, make use of the integrability of the gyroscope to discuss the stability of the three uniform rotations.

5. Suppose that the differential equation in $\mathcal{R}^d$, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, admits a prime integral $A(\mathbf{x})$, i.e. a function $A \in C^\infty(\mathcal{R}^d)$ such that, $\forall\, \mathbf{x} \in \mathcal{R}^d$, $\forall\, t \ge 0$, it is $A(S_t(\mathbf{x})) \equiv A(\mathbf{x})$. Suppose that A has a strict minimum at $\mathbf{x}_0 \in \mathcal{R}^d$. Show that $\mathbf{x}_0$ is a stable stationary point.

6. Use Problem 5 and the conservation of energy to discuss the stationary rotations of the frictionless gyroscope (with $I_1 < I_2 < I_3$) and their stability properties along the following lines. First find the Deprit variables of the uniform stationary rotations (see §4.11, p.317 and p.320) around the inertia axis $\mathbf{i}_k$, $k = 1, 2, 3$. (*Answer:* $K_z, A, A, \gamma, \varphi, \psi + \omega t$ for $\mathbf{i}_3$). Then, using the Deprit Hamiltonian as a prime integral and Problem 5, show that the rotation around the $\mathbf{i}_3$ axis is stable if $I_3 > I_2, I_3$. (*Hint:* Note tat the Deprit Hamiltonian can be written as

$$H = \frac{A^2}{2I_3} + \frac{1}{2}\Big(\frac{\sin^2\psi}{I_1} + \frac{\cos^2\psi}{I_2}\Big)(A^2 - L^2)$$

which has a minimum when $A = L$ if and only if $I_3 > I_2, I_3$.)

7. If the differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ on $\mathcal{R}^d$ is such that there exists a function $A \in C^\infty(U)$, $U \subset \mathcal{R}^d$, which is monotonically non increasing along the motions (i.e. $A(S_t(\mathbf{x})) \le A(\mathbf{x})$, $\forall\, t \ge 0$, $\forall\, \mathbf{x} \in U$, as long as $S_\tau(\mathbf{x}) \in U$, $\forall\, \tau \in [0, t]$, we shall say that A is a monotonic function for the given differential equation in the domain U. If A is monotonically decreasing we call it a "Lyapunov function" for the differential equation. Show that every point where a Lyapunov function has a strict minimum is a stable fixed point.

8. In the context of Problem 7, and under the assumptions of Proposition 6, define

$$A(\mathbf{w}) = \int_0^{+\infty} |S_t(\mathbf{w})|^2 2^{\frac{t}{t_0}}\, dt$$

for $|\mathbf{w}|$ small enough, say, $|\mathbf{w}| < \varrho$. Show that:
(i) A is well defined for all $|\mathbf{w}| < \overline{\varrho}$ if $\overline{\varrho}$ is chosen as in Eq. (5.4.29).
(ii) $A \in C^\infty(U_{\overline{\varrho}})$, where $U_{\overline{\varrho}} = \{\mathbf{w} \,|\, |\mathbf{w}| < \overline{\varrho}\}$.
(iii) A is a Lyapunov function in the sense of Problem 7.
(iv) $2^{\frac{t}{t_0}} A(S_t((\mathbf{w}))$ is monotonic in $t \ge 0$, $\forall\, \mathbf{w} \in U_{\overline{\varrho}}$.
(v) A has a strict minimum at $\mathbf{w} = \mathbf{0}$. This is the "second Lyapunov theorem" (on the existence of a Lyapunov function whenever a stationary point has a stability matrix with eigenvectors with negative real part).

9. Compute the function A of Problem 8 for the linear equation $\dot{\mathbf{x}} = L\mathbf{x}$, supposing that all the eigenvalues of L are pairwise distinct and have a negative real part. Show that A is a positive definite quadratic form in $\mathbf{w}$. (*Answer:* If $\gamma_0 = \frac{1}{2t_0}\log 2 > \frac{\nu \log 2}{2}\log 4A$, with A being the constant introduced in Eq. (5.4.16) and not to be confused with the quadratic form A that we wish to compute, it is $A(\mathbf{w}) = \sum_{i,j=1}^d (\overline{\lambda}_i + \lambda_j + \gamma_0)^{-1} \overline{\alpha_i(\mathbf{w})}\alpha_j(\mathbf{w})$, where $\alpha_i(\mathbf{w})$ is defined as in Eq. (5.4.15).)

10. In the context of Problem 9, show that the ellipsoid $A(\mathbf{w}) = a > 0$ has in $\mathbf{w}$ an outer normal $\mathbf{n}(\mathbf{w})$ such that $\mathbf{n}(\mathbf{w}) \cdot L\mathbf{w} < 0$. (*Hint:* Note that $\mathbf{n}(\mathbf{w}) = \frac{\partial_{\mathbf{w}} A(\mathbf{w})}{|\partial_{\mathbf{w}} A(\mathbf{w})|}$ ($\partial_{\mathbf{w}}$ denoting the gradient); furthermore, by Problem 9 (iii), the derivative of $2^{\frac{t}{2t_0}} A(S_t(\mathbf{w}))$ is non positive:

$$\Big(\frac{A\log 2}{2t_0}+\frac{dA}{dt}\Big)2^{\frac{t}{t_0}}\le 0\ \Rightarrow\ \frac{dA}{dt}\le\frac{-\log 2}{2t_0}A,$$

so if $dA/dt = 0$, it must be that $A = 0$, i.e., $\mathbf{w} = \mathbf{0}$ because A is positive definite. However, $dA/dt = \partial_{\mathbf{w}}(\mathbf{w}\cdot L\mathbf{w})$; hence, $\partial_{\mathbf{w}}(\mathbf{w}\cdot L\mathbf{w}) < 0$ if $\mathbf{w}\neq\mathbf{0}$.)

11. Show that the proof of Proposition 6 can be interpreted as saying that if $A_0(\mathbf{w})$ denotes the Lyapunov function of the linear differential equation $\dot{\mathbf{x}} = L\mathbf{x}$, see Problems 9 and 10, and if $A(\mathbf{w})$ is the Lyapunov function of the differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ with the origin as a fixed point with Lyapunov matrix L, then, assuming that the real part of the eigenvalues of L is negative, $A(\mathbf{w}) = A_0(\mathbf{w}) + O(|\mathbf{w}|^3)$.

12. Consider a one-parameter family of differential equations in $\mathcal{R}^d$: $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\alpha)$, with $\mathbf{x}_0 = \mathbf{0}$ being a stationary point for all values of $\alpha\in(a,b)\subset\mathcal{R}$. Suppose that the Lyapunov matrix of $\mathbf{0}$, $L(\alpha)$ has pairwise distinct eigenvalues, all with real part $\le -\nu < 0$, $\forall\,\alpha\in(a,b)$. Let $\alpha_0\in(a,b)$ and let A_{α_0} be the Lyapunov function of Problem 11, relative to the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\alpha_0)$. Show the existence of $\delta > 0$, $\varepsilon > 0$ such that the neighborhood $V_\delta \stackrel{def}{=} \{\mathbf{x}|\, A_{\alpha_0}(\mathbf{x}) < \delta\}$ has an outer normal $\mathbf{n}(\mathbf{x})$ such that, $\forall\,\mathbf{x}\in\partial V_\delta$ it is $\mathbf{n}(\mathbf{x})\cdot\mathbf{f}(\mathbf{x},\alpha) < 0$, $\forall\,\alpha\in[\alpha_0-\varepsilon,\alpha_0+\varepsilon]$. (*Hint:* First consider the linear case, then Problems 10 and 11).

13. From Problem 12, deduce that V_δ is invariant for the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\alpha)$ for all $\alpha\in[\alpha_0-\varepsilon,\alpha_0+\varepsilon]$. (*Hint:* Suppose the contrary and proceed per absurdum.)

14. Consider a Hamiltonian differential equation in $\mathcal{R}^{2d}$ associated with the Hamiltonian function $H(\mathbf{p},\mathbf{q}) = \frac{1}{2}\mathbf{p}^2 + V(\mathbf{q})$. Let $(\mathbf{0},\mathbf{q}_0)$ be an equilibrium point. Show that its Lyapunov matrix has eigenvalues that can be collected into pairs of opposite value, either both real or both purely imaginary. Furthermore, show that this implies that its stability cannot be settled on the basis of the Lyapunov criterion, while its instability can sometimes be settled on this basis. (*Hint:* Note that the Lyapunov matrix has the structure $L = \begin{pmatrix} A & -B\\ C & D\end{pmatrix}$ where A,B,C,D are the $d\times d$ matrices

$$A = 0,\quad B_{ij} = \frac{\partial^2 V}{\partial q_i\partial q_j},\quad C = 1,\quad D = 0.$$

So if $L\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix}$ with $\mathbf{u},\mathbf{v}\in\mathcal{R}^d$, it must be that $\lambda\mathbf{u} + B\mathbf{v} = \mathbf{0}, \mathbf{u} = \lambda\mathbf{w}$ so that $-\lambda^2\mathbf{v} = B\mathbf{v}$. But B is symmetric so that its eigenvalues are real (see Appendix F), hence ...).

15. Show that the Lyapunov matrix eigenvalues are invariant under regular changes of coordinates $\mathbf{y} = \mathbf{a}(\mathbf{x})$. (*Hint:* If $\boldsymbol{\sigma}$ is defined in the vicinity of the stationary point $\mathbf{x}_0\in\mathcal{R}^d$ for $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and if $J_{ij}(\mathbf{y}) = \frac{\partial\sigma_i(\mathbf{x})}{\partial x_j}$, for $\mathbf{y} = \boldsymbol{\sigma}(\mathbf{x})$, is the Jacobian matrix of the nonsingular change of coordinates, (i.e., such that $\det J\neq 0$), then the differential equation becomes, in $\mathbf{y}$ coordinates, $\dot{\mathbf{y}} = J(\mathbf{y})f(\boldsymbol{\sigma}^{-1}(\mathbf{y}))$, and this implies that the Lyapunov matrix at $\mathbf{y}_0 = \boldsymbol{\sigma}^{-1}(\mathbf{x}_0)$ is $L' = J(\mathbf{y}_0)LJ(\mathbf{y}_0)^{-1}$; hence, $\det(L'-\lambda) = \det(JLJ^{-1}-\lambda) = \det(J(L-\lambda)J^{-1}) = \det(L-\lambda)$.)

16. Let H be a Hamiltonian function describing in some local system of coordinates N point masses in $\mathcal{R}^d$ subject to conservative active forces and constrained by a bilateral ideal constraint to a surface Σ (in the sense of Chapter 3).
Let $(\mathbf{0},\mathbf{x}_0)$, $\mathbf{x}_0\in\Sigma$, be a stationary point. Give arguments (or prove) that the eigenvalues of the Lyapunov matrix for the Hamiltonian equations corresponding to the given stationary point appear in pairs of opposite eigenvalues either both real or both purely imaginary. This is a refinement of Problem 15 extending it to the case of a system ideally constrained to Σ. (*Hint:* In a system of local regular coordinates around $\mathbf{x}_0$ and adapted to Σ, the Lagrangian takes the form [see Eq. (3.11.23), p.215]: $\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^d g(\boldsymbol{\beta})_{ij}\dot\beta_i\dot\beta_j - V(\boldsymbol{\beta})$, with g being

a C^∞ positive-definite matrix function and with V also of class C^∞. So the Hamiltonian is [see Eq. (3.11.25), p.215] $H = \frac{1}{2}\sum_{i,j=1}^{d} g(\boldsymbol{\beta})^{-1}_{ij} p_i p_j + V(\boldsymbol{\beta})$. Hence, the matrix L is $\begin{pmatrix} A & -B \\ C & D \end{pmatrix}$ with

$$A = 0, \quad B_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j}, \quad C = G^{-1}, \quad D = 0,$$

where $G_{ij} = g(\boldsymbol{\beta}_0)_{ij}$ and $\boldsymbol{\beta}_0$ is the point representing $\mathbf{x}_0$ in our system of coordinates. So if $L\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} = \lambda\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix}$, $\mathbf{u},\mathbf{v} \in \mathcal{R}^\ell$, this means $B\mathbf{v} + \lambda\mathbf{u} = \mathbf{0}, G^{-1}\mathbf{u} = \lambda\mathbf{v}$, i.e., $(B + \lambda^2 G)\mathbf{v} = \mathbf{0}$; hence, $0 = \det(B + \lambda^2 G) = \det\left(B + \lambda^2\sqrt{G}\sqrt{G}\right) = \det\left(\sqrt{G}(\sqrt{G^{-1}}B\sqrt{G^{-1}} + \lambda^2)\sqrt{G}\right) = (\det G)\,\det(\sqrt{G^{-1}}B\sqrt{G^{-1}} + \lambda^2)$, see Appendix F for the definition of the square root of a positive-definite matrix). So, since $\sqrt{G^{-1}}B\sqrt{G^{-1}}$ is a symmetric matrix (because G is such, see Appendix F), it follows that λ^2 is real, positive or negative, etc.).

17. Show that Proposition 6 holds if the hypothesis of bounded trajectories is weakened into that of normality or even into no assumption at all; in the latter case, show that global solutions exist for $t \geq 0$ for initial data close enough to $\mathbf{x}_0$. (*Hint:* Simply carefully examine the proof of Proposition 6.)

5.5 Application to the Model of §5.1. The Notion of Vague Attractivity of a Stationary Point

In the case of Eq. (5.1.19), it is easy to compute the Lyapunov matrix relative to the stationary solution $\widehat{\boldsymbol{\omega}}$:

$$L = \begin{pmatrix} \alpha - \lambda_1 - \lambda_2'''\widehat{\omega}_3^2 & -\widehat{\omega}_3 & 0 \\ o_3^3 & \alpha - \lambda_1 - \lambda_2'''\widehat{\omega}_3^2 & 0 \\ 0 & 0 & -\lambda_1 - 3\lambda_2'''\widehat{\omega}_3^2 \end{pmatrix}, \tag{5.5.1}$$

whose eigenvalues are

$$(\alpha - \lambda_1 - \lambda_2'''\widehat{\omega}_3^2) \pm i\widehat{\omega}_3. \tag{5.5.2}$$

Hence, $\widehat{\boldsymbol{\omega}}$ is stable and attractive for some of its neighborhoods not only if $\alpha < \lambda_1$ as already seen in §5.2 and §5.3, but also for $\lambda_1 \leq \alpha < \lambda_1 + \lambda_2'''\widehat{\omega}_3^2$, see Proposition 6, §5.4, p.382. The attractivity of $\widehat{\boldsymbol{\omega}}$ in this interval of variability of α is exponential near $\widehat{\boldsymbol{\omega}}$:

$$\text{if} \quad \alpha - \lambda_1 - 3\lambda_2'''\widehat{\omega}_3^2 < 0 \qquad \text{then} \tag{5.5.3}$$

$$|S_t(\boldsymbol{\omega}) - \widehat{\boldsymbol{\omega}}| \leq 2 \cdot 2^{\frac{t}{t_0}} |\boldsymbol{\omega} - \widehat{\boldsymbol{\omega}}|. \tag{5.5.4}$$

if $|\boldsymbol{\omega} - \widehat{\boldsymbol{\omega}}|$ is small enough; $t_0 > 0$ depends only on the matrix L [see §5.4, comment after Eq. (5.4.30)], and it can be estimated as inversely proportional to $(\lambda_1 + \lambda_2'''\widehat{\omega}_3^2) - \alpha$.

A discussion identical to the one developed in the case $\alpha < \lambda_1$ shows that Eq. (5.5.3) implies that every motion of the gyroscope associated with an evolution $t \to S_t(\mathbf{w})$ like Eq. (5.5.3), for the angular velocity of the comoving

frame, asymptotically tends to become a uniform rotation around the $\mathbf{i}_3$ axis which, in turn, tends to a fixed position in space.

The difference between the cases $\alpha < \lambda_1$, and $\lambda_1 \le \alpha < \lambda_1 + \lambda_2''' \widehat{\omega}_3^2$ lies in the fact that now we can no longer guarantee that $\widehat{\boldsymbol{\omega}}$ is a "global attractor", i.e., with basin of attraction coinciding with $\mathcal{R}^3$. The criterion of Lyapunov has, in fact, only a local character, and thus it can only lead to the recognition of local stability, instability, or attractivity.

Of course, it is of interest to investigate whether or not the attraction basin for $\widehat{\boldsymbol{\omega}}$ is all of $\mathcal{R}^3$, and if not, it would be important to understand where the other attractors for the equation are located. However, this analysis could not be done using general results such as the Lyapunov criterion and we shall not discuss this point in further detail, contenting ourselves with the local information found so far. In any event, it has to be stressed that these kinds of problems are very difficult and very little understood in general.

The motion $\widehat{\boldsymbol{\omega}}$ is no longer stable for $\alpha > \lambda_1 + \lambda_2 \widehat{\omega}_3^2$, not even locally, by the second part of Proposition 6, §5.4. We then inquire about what happens to a solution of Eq. (5.1.19) following an initial datum $\boldsymbol{\omega}$ slightly different from $\widehat{\boldsymbol{\omega}}$ and for $\alpha > a_c = \lambda_1 + \lambda_2 \widehat{\omega}_3^2$, at least for small $\alpha - \alpha_c$.

The first question is whether for α slightly larger than α_c,

$$\alpha_c = \lambda_1 + \lambda_2''' \widehat{\omega}_3^2, \tag{5.5.5}$$

the motion of the data $\boldsymbol{\omega}$ close to $\widehat{\boldsymbol{\omega}}$ departs very much from the motion $\widehat{\boldsymbol{\omega}}$. As we shall see, this question naturally leads to the following interesting notion of "vague attractivity".

5 Definition. *Let $(\mathbf{x}, \alpha) \to \mathbf{f}(\mathbf{x}, \alpha)$ be an $\mathcal{R}^d$-valued $C^\infty(\mathcal{R}^d \times I)$ function with I = open interval, such that the differential equations*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha), \tag{5.5.6}$$

parameterized by $\alpha \in I$, have uniformly bounded trajectories[8] *with respect to $\alpha \in I$ and, furthermore, admit a stationary solution $\mathbf{x}_0 \in \mathcal{R}^d$ such that*

$$\mathbf{f}(\mathbf{x}_0, \alpha) = \mathbf{0}. \tag{5.5.7}$$

$\mathbf{x}_0$ will be called "vaguely attractive" near $\alpha_c \in I$ if there is a neighborhood U of $\mathbf{x}_0$ such that for every $\delta > 0$, one can find $t_\delta > 0$, $\varepsilon_\delta > 0$, $\varrho_\delta > 0$ such that

$$\begin{aligned} S_t^{(\alpha)}(U) &\subset \Gamma(\delta), \qquad \forall\, t \ge t_\delta,\ \forall\, \alpha \in (\alpha_c - \varepsilon_\delta, \alpha_c + \varepsilon_\delta), \\ S_t^{(\alpha)}(\Gamma(\delta)) &\subset \Gamma(\varrho_\delta), \qquad \forall\, t \ge t_\delta,\ \forall\, \alpha \in (\alpha_c - \varepsilon_\delta, \alpha_c + \varepsilon_\delta), \end{aligned} \tag{5.5.8}$$

with $\varrho_\delta \xrightarrow[\delta \to 0]{} 0$. Here $S_t^{(\alpha)}$ is the solution flow for Eq. (5.5.5) and $\Gamma(\delta)$ = cube with side 2δ centered around $\mathbf{x}_0$.

[8] If $S_t^{(\alpha)}$ denotes the flow generated by Eq. (5.5.5), this means that the bound on the trajectory of $\mathbf{x} \in \mathcal{R}^d$, $|S_t^{(\alpha)}(\mathbf{x})| \le \mu(|\mathbf{x}|)$ holds and μ is continuous and α-independent.

Observations.
(1) In other words, $\mathbf{x}_0$ is vaguely attractive near α_c, if there is a neighborhood U which is a basin of attraction for an attractor containing $\mathbf{x}_0$ and having a diameter smaller than any arbitrarily prefixed length $\delta > 0$ for all α's close enough to α_c. Furthermore, this attractor, contained in $\Gamma(\delta)$, "uniformly attracts" the points of U and has a "weak stability", as expressed more precisely by the first and second of Eqs. (5.5.7), respectively.
Note that for $\alpha = \alpha_c$, the point $\mathbf{x}_0$ must be attractive for the points in U. In fact $\mathbf{x}_0$ is vaguely attractive for α near α_c if and only if it is stable and attractive for Eq. (5.5.6) with $\alpha = \alpha_c$.
(2) One can also say that $\mathbf{x}_0$ is vaguely attractive near α_c if it is the attractor of a neighborhood U of $\mathbf{x}_0$, for $\alpha = \alpha_c$ while for α close to α_c it still attracts the points of U not too close to $\mathbf{x}_0$. The "attractivity away from $\mathbf{x}_0$ is uniform in α" near α_c.
(3) If in α_c the Lyapunov matrix $L(\alpha)$ for Eq. (5.5.5) relative to $\mathbf{x}_0$ has eigenvalues with a negative real part, it follows from the arguments of the proof of Proposition 6, §5.4, that $\mathbf{x}_0$ is vaguely attractive near α_c. Actually, the set U can be taken such that for some $\varepsilon_0 > 0$ it is $S_t^{(\alpha)} U \subset U, \forall\, |\alpha - \alpha_c| < \varepsilon_0, \forall\, t \geq 0$, (this follows from the Problems 12 and 13, §5.4, p.388.)
(4) Hence, the vague-attractivity notion is interesting only when $L(\alpha_c)$ has some eigenvalues with a vanishing real part.
(5) All the upcoming examples of vague attractivity will have the property that U can be chosen to fulfill Eq. (5.5.8) for all $t \geq 0$. It seems not impossible that the neighborhood U of vague attractivity could always be chosen in this way.
(6) The condition that $\Gamma(\delta)$ be a cube with side 2δ centered at $\mathbf{x}_0$ could be equivalently replaced by the requirement that $\Gamma(\delta)$ be a family of neighborhoods of $\mathbf{x}_0$ with diameter tending to zero with δ.
(7) The assumption that $\mathbf{x}_0$ should be α independent is only apparently more restrictive than the natural assumption of the existence of a stationary solution $\mathbf{x}^{(\alpha)}$ depending in a C^∞-regular way on α. With a change of coordinates, one can always reduce the stability theory, of such a stationary point, to that relative to the case when $\mathbf{x}^{(\alpha)} = \mathbf{0}$.
(8) If $\mathbf{x}_0$ is replaced in the above definition by an invariant set A and $\Gamma_A(\delta) = \{$set of the points at distance $< \delta$ from $A\}$, one defines the notion of a "vague attractor".
This notion could be extended to the case when A depends on α, although not as straightforwardly and as unambiguously, as in the case $A = \{\mathbf{x}^{(\alpha)}\}$ discussed in Observation (7).
(9) Last but not least, the vague attractivity of $\mathbf{x}_0$ is a notion invariant under changes of coordinates; it is also invariant under changes of the equation itself (i.e., of the function $\mathbf{f}(\mathbf{x}, \alpha)$, for $\mathbf{x}$ outside some neighborhood of $\mathbf{x}_0$). Vague attractivity is an "intrinsic local property" of Eq. (5.5.5) near $\mathbf{x}_0$ and α_c.

The above facts play an important role in the formulation of simple vague attractivity criteria which show that it is a property that can be inferred from the knowledge of the $\mathbf{x}$ derivatives of $\mathbf{f}(\mathbf{x},\alpha_c)$ in $\mathbf{x}_0$ of order not exceeding 3. To illustrate this important fact and to provide, in this way, some simple vague-attractivity criteria, it is convenient to introduce the notion of "normal form" of a differential equation near a stationary solution.

6 Definition. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\alpha)$, $\mathbf{f} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$, be a differential equation in $\mathcal{R}^d$ with uniformly bounded trajectories (see footnote 8 to p.383) and with $\mathbf{x}_0 = \mathbf{0}$ as a stationary point $\forall\,\alpha \in I = (a,b)$.*
Let $L(\alpha)$ be the stability matrix at $\mathbf{x}_0 = 0$ and suppose that $\lambda_1(\alpha), \overline{\lambda_1(\alpha)}, \ldots, \overline{\lambda_p(\alpha)}, \lambda'_1(\alpha), \ldots, \lambda'_q(\alpha)$ are $2p+q$ of its eigenvalues, the first $2p$ being arranged into complex-conjugate non real pairs and the last q being real.
We say that the differential equation has, for $\alpha \in I$, a "normal form" with respect to the mentioned eigenvalues of $L(\alpha)$ if, writing the coordinates of $\mathbf{x}$ as $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(p)}, y^{(1)}, \ldots, y^{(q)}, \mathbf{z})$ with $\mathbf{x}^{(j)} \in \mathcal{R}^2$, $j = 1, \ldots, p$, $y^{(i)} \in \mathcal{R}$, $i = 1, \ldots, q$, and $\mathbf{z} \in \mathcal{R}^{d-2p-q}$, the equation has the form, $\forall\,\alpha \in I$,

$$\begin{aligned}
\dot{\mathbf{x}}_1^{(j)} =& (\mathcal{R}\mathrm{e}\,\lambda_1(\alpha))\,x_1^{(j)} - (\mathcal{I}\mathrm{m}\,\lambda_1(\alpha))\,x_2^{(j)} \\
&+ N_1^{(j)}(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}, y^{(1)}, \ldots, y^{(p)}, \mathbf{z}, \alpha), \\
\dot{\mathbf{x}}_2^{(j)} =& (\mathcal{I}\mathrm{m}\,\lambda_1(\alpha))\,x_1^{(j)} - (\mathcal{R}\mathrm{e}\,\lambda_1(\alpha))\,x_2^{(j)} \\
&+ N_2^{(j)}(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}, y^{(1)}, \ldots, y^{(p)}, \mathbf{z}, \alpha), \\
\dot{y}^{(h)} =& \lambda'_h(\alpha) y^{(h)} + M^{(h)}(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}, y^{(1)}, \ldots, y^{(p)}, \mathbf{z}, \alpha), \\
\dot{\mathbf{z}} =& \widetilde{L}(\alpha) + \widetilde{\mathbf{P}}(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}, y^{(1)}, \ldots, y^{(p)}, \mathbf{z}, \alpha),
\end{aligned} \tag{5.5.9}$$

$j = 1, \ldots, p$, $h = 1, \ldots, q$, $\widetilde{L}(\alpha)$ being a $(d-2p-q) \times (d-2p-q)$ matrix with C^∞ entries (as functions of α) and $\mathbf{N}^{(j)}, M^{(h)}, \widetilde{\mathbf{P}}$ being C^∞ functions of their arguments with the extra property that $\mathbf{N}^{(j)}, M^{(h)}$ have a zero of third order at the origin in the $\mathbf{x}, y, \mathbf{z}$ variables, for all $\alpha \in I$, while $\widetilde{\mathbf{P}}$ has a second-order zero, at least, at the origin (in the same variables).

Observation. If $p = 0$ or $q = 0$ or $d = 2p + q$, the above definition makes sense in an obvious way by deleting parts of Eq. (5.5.9).

Vague attractivity near α_c may be easily discussed once the equation is in normal form with respect to the eigenvalues of $L(\alpha)$ whose real part vanishes for $\alpha = \alpha_c$, In general, the equations that one wishes to study will not have normal form, but they may acquire such a form after a change of variables. This is as suitable for vague-attractivity analysis, by Observation (9), p.391.

For instance, Eq. (5.1.19) does not have normal form near α_c with respect to the two complex eigenvalues of $L(\alpha)$. Therefore, before discussing a vague-attractivity criterion, it is convenient to remark that there is a simple and rather weak sufficient condition for the existence of a system of coordinates where the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\alpha)$ assumes normal form.

7 Proposition. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha), \mathbf{f} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$, be a differential equation parameterized by α and with uniformly bounded trajectories as $\alpha \in I = (a, b)$. Suppose that $\mathbf{f}(\mathbf{0}, \alpha) = \mathbf{0}$ and let $L(\alpha)$ be the stability matrix of $\mathbf{0}$.*
Suppose that for $\alpha \in I$, $L(\alpha)$ has d pairwise-distinct eigenvalues $\Lambda_1(\alpha), \ldots, \Lambda_d(\alpha)$, among which $2p$ are non real; write them as $\lambda_1(\alpha), \overline{\lambda_1(\alpha)}, \ldots, \overline{\lambda_p(\alpha)}$, $\lambda'_1(\alpha), \ldots, \lambda'_q(\alpha)$, $(2p + q = d)$, and arrange them so that the functions $\alpha \to \lambda_i(\alpha)$ are C^∞-functions of α, for $\alpha \in I$.[9]
(i) There is a (global) coordinate system on $\mathcal{R}^d \times I$: $(\mathcal{R}^d \times I, \boldsymbol{\Xi})$, with basis $\mathcal{R}^d \times I$, denoted $(\mathbf{x}, \alpha) = \boldsymbol{\Xi}(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(p)}, \boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(q)}, \alpha')$ with $\alpha' = \alpha, \boldsymbol{\xi}^{(j)} \in \mathcal{R}^2, \eta^{(h)} \in \mathcal{R}$, such that in the new coordinates, the equations takes the form, $(j = 1, \ldots, p;\ h = 1, \ldots, q)$,

$$\begin{aligned}
\dot{\xi}_1^{(j)} &= ((\mathcal{R}\mathrm{e}\,\lambda_1(\alpha))\,\xi_1^{(j)} - (\mathcal{I}\mathrm{m}\,\lambda_1(\alpha))\,\xi_2^{(j)}) + F_1^{(j)}(\boldsymbol{\xi}^{(1)}, \ldots, \alpha),\\
\dot{\xi}_2^{(j)} &= ((\mathcal{I}\mathrm{m}\,\lambda_1(\alpha))\,\xi_1^{(j)} - (\mathcal{R}\mathrm{e}\,\lambda_1(\alpha))\,\xi_2^{(j)}) + F_2^{(j)}(\boldsymbol{\xi}^{(1)}, \ldots, \alpha),\\
\dot{\eta}^{(h)} &= \lambda'_h\,\eta^{(h)} + F^{(h)}(\boldsymbol{\xi}^{(1)}, \ldots, \alpha),
\end{aligned} \tag{5.5.10}$$

where $F_1^{(j)}, F_2^{(j)}, F^{(h)}$ are in $C^\infty(\mathcal{R}^d \times I)$ and have a second-order zero at the origin in the variables $\boldsymbol{\xi}, \boldsymbol{\eta}$ for each $\alpha \in I$.
(ii) If $\Lambda_k(\alpha_c) \neq \Lambda_k(\alpha_c) + \Lambda_\ell(\alpha_c)$, $k, h, \ell = 1, \ldots, d$, and if the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ has already the form of Eq. (5.5.10) for $a \in I$, there is a coordinate system on a suitable neighborhood $U \times J$ of $(\mathbf{0}, \alpha_c)$, $(U \times J, \boldsymbol{\Xi})$, such that in the new coordinates, Eq. (5.5.10) takes normal form with respect to all the eigenvalues of $L(\alpha)$. Calling $(\boldsymbol{\beta}, \alpha')$ the new coordinates, the transformation $\boldsymbol{\Xi}$ can be chosen as

$$\beta_j = x_j - \sum_{k,\ell=1}^{d} S_{jk\ell}(\alpha)\, x_k\, x_\ell, \qquad \alpha' = \alpha \tag{5.5.11}$$

with $S_{jk\ell} = S_{j\ell k} \in C^\infty(J)$, a "quadratic change of coordinates"; its inverse will (therefore[10]*) have the form*

$$x_j = \beta_j + \sum_{k,\ell=1}^{d} S_{jk\ell}(\alpha)\, \beta_k\, \beta_\ell + G_j(\boldsymbol{\beta}, \alpha), \qquad \alpha' = \alpha, \tag{5.5.12}$$

where $G_j \in C^\infty(\boldsymbol{\Xi}(U \times J))$ has a third-order zero at $\boldsymbol{\beta} = \mathbf{0}$, $\forall\, \alpha \in J$.

Observations.
(1) Note that by defining, for $(j = 1, \ldots, p;\ h = 1, \ldots, q)$

[9] Since the eigenvalues are supposed to be pairwise distinct and they are roots of a d-th order polynomial, this is possible and it follows from general results in Algebra.
[10] By the implicit function theorem, (see Appendix G).

$$\begin{aligned}
&z^{(j)} = \xi_1^{(j)} + i\,\xi_2^{(j)}, \qquad \overline{z}^{(j)} = \xi_1^{(j)} - i\,\xi_2^{(j)},\\
&N^{(j)}(z^{(1)}, \overline{z}^{(1)}, \dots, z^{(p)}, \overline{z}^{(p)}, \eta^{(1)}, \dots, \eta^{(q)}, \alpha)\\
&= F_1^{(j)}(\boldsymbol{\xi}^{(1)}, \dots) + i\,F_2^{(j)}(\boldsymbol{\xi}^{(1)}, \dots),\\
&M^{(h)}(z^{(1)}, \overline{z}^{(1)}, \dots, z^{(p)}, \overline{z}^{(p)}, \eta^{(1)}, \dots, \eta^{(q)}, \alpha) = F^{(h)}(\boldsymbol{\xi}^{(1)}, \dots)
\end{aligned} \tag{5.5.13}$$

Eq. (5.5.10) assumes the more symmetric form

$$\begin{aligned}
\dot{z}^{(j)} =& \lambda_j(\alpha) z^{(j)} + N^{(j)}(z^{(1)}, \overline{z}^{(1)}, \dots, \alpha), \qquad j = 1, \dots, p,\\
\dot{\eta}^{(h)} =& \lambda'_h(\alpha)\eta^{(h)} + M^{(h)}(z^{(1)}, \overline{z}^{(1)}, \dots, \alpha), \qquad h = 1, \dots, q,
\end{aligned} \tag{5.5.14}$$

(2) If the eigenvalues $\lambda_1(\alpha_c), \overline{\lambda_1(\alpha_c)}$ are non degenerate and $\mathcal{R}\mathrm{e}\,\lambda_1(\alpha_c) = 0, \mathcal{I}\mathrm{m}\,\lambda_1(\alpha_c) \neq 0$, it follows from (ii) that it will be possible to put the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ into normal form with respect to $\lambda_1(\alpha_c), \overline{\lambda_1(\alpha_c)}$ in the sense of Definition 6 above. This is obvious if $\Lambda_k(\alpha_c) \neq \Lambda_h(\alpha_c) + \Lambda_\ell(\alpha_c)$, $\forall\, k, h, \ell$, but it is also generally true as a consequence of (ii) (see below).
Suppose, in fact, that the equation has already the form of Eq. (5.5.10). We then perform the quadratic change of coordinates that would put into normal form, (with respect to all the eigenvalues), the equation obtained from Eq. (5.5.10) by replacing the eigenvalues $\lambda_1(\alpha), \overline{\lambda_1(\alpha)}, \dots, \overline{\lambda_p(\alpha)}$, $\lambda'_1(\alpha), \dots, \lambda'_q(\alpha)$, by $\widetilde{\Lambda}_1(\alpha), \dots, \widetilde{\Lambda}_d(\alpha)) = \lambda_1(\alpha), \overline{\lambda_1(\alpha)}, \lambda_2(\alpha) + \varepsilon_2, \overline{\lambda_2(\alpha)} + \overline{\varepsilon}_2, \dots,$, $\lambda'_1(\alpha) + \varepsilon'_1, \dots,\ \lambda'_q(\alpha) + \varepsilon'_q$, where $\varepsilon_2, \dots, \varepsilon'_1, \dots$ are chosen so that the condition $\widetilde{\Lambda}_k(\alpha) + \widetilde{\Lambda}_h(\alpha) \neq \widetilde{\Lambda}_\ell(\alpha)$, $\forall\, k, h, \ell$ is fulfilled (and the ε'_h are real).
Taking into account the quadratic nature of the maps of Eqs. (5.5.11) and (5.5.12), it is clear that the original equation will take, in the new coordinates, normal form with respect to the only two eigenvalues which have not been modified, i.e., $\lambda_1(\alpha), \overline{\lambda_1(\alpha)}$. If the equation does not have the form of Eq. (5.5.10), but $\lambda_1(\alpha), \overline{\lambda_1(\alpha)}$ are non degenerate, one can apply a similar argument.
(3) From the proof, it appears that the normal-form coordinates (for all the eigenvalues or, via the previous observations (2) suitably adapted, for some of them) can sometimes be found even when the "non resonance condition" on the eigenvalues [in (ii) above] is not fulfilled, provided the equation verifies additional properties. Such conditions can be explicitly stated by requiring that Eq. (5.5.22) below be solvable. In the problems 16 and 17 at the end of this section, we give some examples of explicit use of this remark.

Proof. To find $\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(p)}, \eta^{(1)}, \dots, \eta^{(q)}$ coordinates, consider the eigenvectors $(\mathbf{w}^{(1)}(\alpha), \dots, \mathbf{w}^{(d)}(\alpha)) \equiv (\mathbf{v}^{(1)}(\alpha), \overline{\mathbf{v}^{(1)}(\alpha)}, \dots, \mathbf{v}^{(p)}(\alpha), \overline{\mathbf{v}^{(p)}(\alpha)}, \mathbf{v}'^{(1)}(\alpha), \dots, \mathbf{v}'^{(q)}(\alpha))$, of $L(\alpha)$ associated with the eigenvalues $(\lambda_1(\alpha), \overline{\lambda_1(\alpha)}, \dots, \overline{\lambda_p(\alpha)}$, $\lambda'_1(\alpha), \dots,\ \lambda'_q(\alpha))$, $\alpha \in I$, respectively. At fixed $\alpha \in I$, such vectors are linearly independent, by the assumption of distinct eigenvalues.

We may and shall assume that the above eigenvectors are C^∞ functions of α. Since they form a basis in $\mathcal{C}^d$, any $\mathbf{x} \in \mathcal{R}^d$ can be written, defining $\zeta^{(j)} \stackrel{def}{=} \xi_1^{(j)} + i\,\xi_2^{(j)}$, as

$$\mathbf{x} = \sum_{j=1}^{p} \left[\zeta^{(j)}\mathbf{v}^{(j)}(\alpha) + \overline{\zeta}^{(j)}\overline{\mathbf{v}^{(j)}(\alpha)}\right] + \sum_{h=1}^{q} \eta^{(h)}\,\mathbf{v}'^{(h)}(\alpha). \tag{5.5.15}$$

and, remarking that $\mathbf{v}^{(j)}(\alpha), \mathbf{v}'^{(j)}(\alpha)$ are eigenvectors of $L(\alpha)$, it is immediate to check that in the $(\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(p)}, \eta^{(1)}, \ldots, \eta^{(q)})$ coordinates, the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ takes the form of Eq. (5.5.10).

To prove (ii), write Eq. (5.5.10) as

$$\dot{x}_j = f_j(\mathbf{x}, \alpha) = \sum_{k=1}^{d} L_{jk}(\alpha)x_k + \sum_{k,\ell}^{d} F_{jk\ell}x_h x_\ell \tag{5.5.16}$$

with $F_{jk\ell} = F_{j\ell k} \in C^\infty(\mathcal{R}^d \times J)$, $j = 1, \ldots, d$.

Performing the change of coordinates in Eqs. (5.5.11) and (5.5.12), after some algebra Eq. (5.5.16) becomes, in the new coordinates $\boldsymbol{\beta}$,

$$\begin{aligned}\dot{\beta}_j =& \dot{x}_j - 2\sum_{k,\ell} S_{jk\ell}(\alpha)x_k\dot{x}_\ell = \sum_{k=1}^{d} L_{jk}(\alpha)\beta_k + \sum_{h,\ell=1}^{d}\Big\{\sum_{k=1}^{d}\big(L_{jk}S_{kh\ell}(\alpha) \\ &- S_{jhk}(\alpha)L_{k\ell} - S_{j\ell k}(\alpha)L_{kh}(\alpha)\ + F_{jh\ell}(\mathbf{0}, \alpha)\big)\Big\}\,\beta_h\,\beta_\ell + \overline{G}_j(\boldsymbol{\beta}, \alpha)\end{aligned} \tag{5.5.17}$$

where $\overline{G}_j$ has a third-order zero at $\boldsymbol{\beta} = \mathbf{0}$, $\forall\,\alpha \in J$.

Therefore, if there is a solution $S_{jhk}(\alpha)$ to the linear system of $d\frac{d(d+1)}{2}$ equations in $d\frac{d(d+1)}{2}$ unknowns (recall that $S_{jh\ell} = S_{j\ell k}$, $F_{jh\ell} = F_{j\ell k}$) described for $j, h, \ell = 1, \ldots, d$ by

$$\sum_{k=1}^{d}\big(L_{jk}S_{kh\ell}(\alpha) - S_{jhk}(\alpha)L_{k\ell} - S_{j\ell k}(\alpha)L_{kh}(\alpha)\big) + F_{jh\ell}(\mathbf{0}, \alpha) = 0 \tag{5.5.18}$$

and if the solution S_{jhk} depends on α in a C^∞ way for α near α_c, then Proposition 7 will have been proved.

Define a matrix $W(\alpha)$ in terms of the eigenvectors of $L(\alpha)$, $\mathbf{w}^{(1)}(\alpha), \ldots,$ $\mathbf{w}^{(d)}(\alpha)$, as

$$W(\alpha)_{hk} \stackrel{def}{=} w_h^{(k)}(\alpha), \qquad h, k = 1, \ldots, d. \tag{5.5.19}$$

The linear independence of the eigenvectors $\mathbf{w}^{(i)}$ implies that $\det W(\alpha) \neq 0$, $\forall\,\alpha \in I$, so that $W(\alpha)^{-1}$ exists and is a C^∞-matrix function of $\alpha \in I$, and if a matrix $\Lambda(\alpha)$ is defined as $\Lambda(\alpha)_{hk} = \Lambda_h(\alpha)\delta_{hk}$, it is

$$L(\alpha) \;=\; W(\alpha)\, \Lambda(\alpha)\, W(\alpha)^{-1} \tag{5.5.20}$$

(see Appendix F for details on this relation between a matrix, its eigenvalues, and its eigenvectors). Inserting Eq. (5.5.20) into Eq. (5.5.18) one finds

$$\sum_{k,d=1}^{d} \big(W(\alpha)_{js}\Lambda_s(\alpha)W(\alpha)^{-1}_{sk}S_{kh\ell}(\alpha) - S_{jhk}(\alpha)W(\alpha)_{ks}\Lambda_s(\alpha)W(\alpha)^{-1}_{k\ell} \\ - S_{j\ell k}(\alpha)W(\alpha)_{ks}\Lambda_s(\alpha)W(\alpha)^{-1}_{sh}(\alpha)\big) \;+ F_{jh\ell}(\mathbf{0},\alpha)$$

and multiplying both sides by $W(\alpha)^{-1}_{rj}W(\alpha)_{\ell p}W(\alpha)_{hq}$ summing over j,h,ℓ, and setting

$$\begin{aligned} \sigma_{spq}(\alpha) &= \sum_{j,k.\ell} W(\alpha)^{-1}_{sk}S_{kh\ell}(\alpha)W(\alpha)_{hq}W(\alpha)_{\ell p} \\ \varphi_{spq}(\alpha) &= \sum_{j,k.\ell} W(\alpha)^{-1}_{sk}F_{kh\ell}(\mathbf{0},\alpha)W(\alpha)_{hq}W(\alpha)_{\ell p} \end{aligned} \tag{5.5.21}$$

one finds that Eq. (5.5.18) becomes

$$\varphi_{spq} + (\Lambda_s(\alpha) - \Lambda_p(\alpha) - \Lambda_q(\alpha))\sigma_{spq}(\alpha) = 0 \tag{5.5.22}$$

which can certainly be solved uniquely for σ and via Eq. (5.5.21) yields a C^∞ solution to Eq. (5.5.18), $\forall\, \alpha \in J$. This solution is real because Eq. (5.5.18) is a linear equation with real coefficients and real known terms. mbe

Observation. Note that in the above proof, the determination of the change of coordinates leading to the form of Eq. (5.5.10) only involves the matrix $L(\alpha)$, i.e., the first $\mathbf{x}$ derivatives at the origin of $\mathbf{f}(\mathbf{x},\alpha)$. The definition of $S_{jk\ell}(\alpha)$, i.e., of the coordinates putting the equation into the normal form, only involves $L(\alpha)$ and $F_{jk\ell}(\mathbf{0},\alpha)$, i.e., the first and second derivatives of $\mathbf{f}(\mathbf{x},\alpha)$ at $\mathbf{0}$.

It is now possible to discuss a simple vague-attractivity criterion.

8 Proposition. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\alpha)$, $\mathbf{f} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$, be a differential equation parameterized by α, with uniformly bounded trajectories as $\alpha \in I = (a,b)$, (see footnote 8) and such that $\mathbf{f}(\mathbf{0},\alpha) = \mathbf{0}$, $\forall\, \alpha \in I$.*
Suppose that for $\alpha = \alpha_c$, the stability matrix of the origin, $L(\alpha_c)$, has one pair of conjugate imaginary eigenvalues $\lambda_1(\alpha_c), \overline{\lambda_1(\alpha_c)} \neq 0$, while all the other $d-2$ eigenvalues have negative real parts.
Also suppose that the equation has normal form with respect to λ_1, λ_2 near $\alpha_c \in I$, and write the differential equations for the first two components of $\mathbf{x}$, x_1 and x_2, as

$$\begin{aligned} \dot{x}_1 =& (\mathcal{R}\mathrm{e}\,\lambda_1(\alpha)\,x_1 - (\mathcal{I}\mathrm{m}\,\lambda_1(\alpha))\,x_2) + N_1(x_1,x_2,\mathbf{y},\alpha), \\ \dot{x}_2 =& (\mathcal{I}\mathrm{m}\,\lambda_1(\alpha)\,x_1 + (\mathcal{R}\mathrm{e}\,\lambda_1(\alpha))\,x_2) + N_2(x_1,x_2,\mathbf{y},\alpha), \end{aligned} \tag{5.5.23}$$

with $N_1, N_2 \in C^\infty(\mathcal{R}^d \times \mathcal{R})$ *and having a third-order zero at the origin* $x_1 = x_2 = 0, \mathbf{y} = \mathbf{0}$, *for all* $\alpha \in I$, *having denoted* $\mathbf{y}$ *the last* $d-2$ *coordinates of* $\mathbf{x}$. *If* $x_1 + ix_2 \stackrel{def}{=} \varrho e^{i\theta}$ *define*

$$\gamma_\alpha = \lim_{\varrho \to 0} \lim_{\mathbf{y} \to \mathbf{0}} \frac{x_1 N_1 + x_2 N_2}{(x_1^2 + x_2^2)^2} \tag{5.5.24}$$

for $\alpha \in I$. *Then the origin is vaguely attractive near* α_c *if*

$$\overline{\gamma} = \frac{1}{2\pi} \int_0^{2\pi} \gamma_{\alpha_c}(\theta)\, d\theta < 0, \tag{5.5.25}$$

while if $\overline{\gamma} > 0$ *is is not vaguely attractive.*
The same conclusions can be drawn under the sole assumption that the differential equation takes the form of Eq. (5.5.23) without requiring that N_1 *and* N_2 *be of third order, but only requiring the existence of the limit Eq. (5.5.24), i.e., only requiring that* $x_1 N_1 + x_2 N_2$ *be of fourth order.*

Observations.
(1) As already remarked, the assumption on the normality of the equation with respect to $\lambda_1(\alpha), \overline{\lambda_1(\alpha)}$ is not really restrictive if (as assumed above) $\mathcal{I}\mathrm{m}\, \lambda_1(\alpha_c) \neq 0$ and if all the remaining eigenvalues have a negative real part. In fact, one can always change coordinates and put the equation in this form (see observation (2), p.394, to Proposition 7).
(2) The number $\overline{\gamma}$ can, in principle, be computed in any system of coordinates in terms of the derivatives of first order, second order, and third order of $\mathbf{f}(\mathbf{x}, \alpha_c)$ at $\mathbf{x} = \mathbf{0}$, with respect to the $\mathbf{x}$ coordinates. However, this calculation may be very long in practical cases. For the computation of $\overline{\gamma}$, it is more practical to first reduce the equation to the form of Eq (5.5.23) using observation (2) to Proposition 7, p.394, and then to compute $\overline{\gamma}$ via Eq. (5.5.25).
(3) A similar criterion holds if the equation has one real eigenvalue $\lambda'(\alpha)$ vanishing at α_c while all the others remain with negative real part near α_c if $\mathbf{x} = (x_1, \mathbf{y})$ and assuming

$$\dot{x}_1 = \lambda'(\alpha) x_1 + N_1(x_1, \mathbf{y}, \alpha) \tag{5.5.26}$$

with N_1 having a zero of third order at $x_1 = 0, \mathbf{y} = \mathbf{0}$, $\forall\, \alpha \in I$, then a vague-attractivity criterion is that $\overline{\gamma} = \lim_{x_1 \to 0} \frac{x_1 N_1(x_1, \mathbf{0}, \alpha_c)}{x_1^4} < 0$.
However, the above normal-form assumption, i.e., the assumption that N should be of third order, is now restrictive. Sometimes it might be impossible to find coordinates in which the equation for x_1 takes the form of Eq. (5.5.26).

PROOF. For simplicity, we suppose that the only non real eigenvalues of $L(\alpha)$ are $\lambda(\alpha) \equiv \lambda_1(\alpha) = \sigma(\alpha) + i\,\mu(\alpha) = \overline{\lambda_2(\alpha)}$ for α near α_c; we also suppose that the other eigenvalues are pairwise distinct and $\mu(\alpha) > 0$.

Let $\nu > 0, \alpha > 0$ be such that $\lambda_1'(\alpha), \ldots, \lambda_{d-2}'(\alpha) \leq -\nu < 0, \quad \mu(\alpha) > \nu, \quad \forall\, \alpha \in (\alpha_c - a, \alpha_c + a)$. We may and shall suppose that the equation takes the form

$$\begin{aligned}\dot{x}_1 =& \sigma(\alpha)x_1 - \mu(\alpha)x_2 + N_1(x_1, x_2, \mathbf{y}, \alpha),\\ \dot{x}_2 =& \mu(\alpha)x_1 + \sigma(\alpha)x_2 + N_2(x_1, x_2, \mathbf{y}, \alpha),\\ \dot{y}_j =& \lambda_j'(\alpha)y_j + \widetilde{N}_j(x_1, x_2, \mathbf{y}, \alpha),\end{aligned} \tag{5.5.27}$$

$j = 1, \dots, d-2$, with N_1, N_2 having a third-order zero at $x_1 = 0, x_2 = 0, y = 0$, $\forall\, \alpha \in (a_c - a, a_c + a)$, while N_i has at least a second-order zero at the same point, $\forall\, \alpha \in (\alpha_c - a, \alpha_c + a)$, see Proposition 7 (i), p.393.

By the Lagrange-Taylor theorem, see Appendix B, we can write, for

$$\begin{aligned}N_j(x_1, x_2, \mathbf{y}, \alpha) = &\sum_{h,k,\ell=1}^{2} \overline{N}_{jhk\ell}(\alpha)x_h x_k x_\ell\\ + &\sum_{h,k=1}^{2}\sum_{\ell=1}^{d-2} \overline{N}'_{jhk\ell}(\alpha)x_h x_k y_\ell + \sum_{h}^{2}\sum_{k,\ell=1}^{d-2} \overline{N}''_{jhk\ell}(\alpha)x_h y_k y_\ell\\ + &\sum_{h,k,\ell=1}^{d-2} \overline{N}'''_{jhk\ell}(\alpha)y_h y_k y_\ell + \widehat{N}_j(x_1, x_2, \mathbf{y}, \alpha),\end{aligned} \tag{5.5.28}$$

where $\overline{N}, \overline{N}', \overline{N}'', \overline{N}'''$ are C^∞ functions of $\alpha \in (\alpha_c - a, \alpha_c + a)$ and $\widehat{\mathbf{N}}$ is a C^∞ function of its arguments, for $\alpha \in (\alpha_c - a, \alpha_c + a)$, and it has a fourth order zero at the origin in the $x_1, x_2, \mathbf{y}$ variables, $\forall\, \alpha \in (\alpha_c - a, \alpha_c + a)$.

For all $\alpha \in (\alpha_c - a, \alpha_c + a)$, if $x_1 + i\, x_2 \stackrel{def}{=} \varrho e^{i\theta}$, it is

$$\gamma_\alpha(\theta) = \sum_{j,h,k,\ell=1}^{2} \overline{N}_{jhk\ell}(\alpha)\frac{x_j x_k x_h x_\ell}{(x_1^2 + x_2^2)^2}. \tag{5.5.29}$$

To continue, first assume that $\gamma_\alpha(\theta) \equiv \overline{g}_\alpha < 0$, $\forall\, \alpha \in (\alpha_c - a, \alpha_c + a)$, i.e., suppose that $\gamma_\alpha(\theta)$ is θ-independent. This severe restriction will be later removed. Multiply the Eqs. (5.5.27) by x_1, x_2, y_j, respectively, and sum the first two and, separately, the last $d-2$ to find, setting $\varrho \stackrel{def}{=} \sqrt{x_1^2 + x_2^2}$,

$$\frac{1}{2}\frac{d\varrho^2}{dt} = \sigma(\alpha)\varrho^2 + \overline{\gamma}\varrho^4 + D_4, \qquad \frac{1}{2}\frac{dy^2}{dt} \le -\nu y^2 + D_3, \tag{5.5.30}$$

where D_3 and D_4 are C^∞ functions of $x_1, x_2, \mathbf{y}$ and of $\alpha \in (\alpha_c - a, \alpha_c + a)$ such that $\exists\, C_1, C_2 > 0$ which, for $\varrho, \mathbf{y}$ near zero, verify

$$|D_4| \le C_1\,(|\mathbf{y}| + \varrho^2)\,(\varrho + |\mathbf{y}|)^3, \qquad |D_3| \le C_2\,(\varrho + |\mathbf{y}|)^3. \tag{5.5.31}$$

Let $\beta, \delta_0 > 0$ be such that $\beta(1+\beta)^3 C_1 \le \frac{1}{2}|\overline{\gamma}|$, $\nu\beta^2\delta_0^2 > \frac{1}{2}C_2(1+\beta)^3\delta_0^3$, and let δ_0 be so small that for all $\delta \le \delta_0$ Eqs. (5.5.31) hold in

$$\Gamma(\delta) = \{x_1, x_2, \mathbf{y} \,|\, \varrho < \delta,\, |\mathbf{y}| < \delta\beta\}. \tag{5.5.32}$$

Then we see that for $(x_1, x_2, \mathbf{y}) \in \partial\Gamma(\delta)$, $\delta \le \delta_0$, it is

$$\begin{aligned} \frac{1}{2}\frac{d\varrho^2}{dt} &\le \sigma(\alpha)\delta^2 + \frac{\overline{\nu}}{2}\delta^4, \qquad \text{if } \varrho = \delta,\\ \frac{1}{2}\frac{d\mathbf{y}^2}{dt} &\le -\frac{\nu}{2}\beta^2\delta^2, \qquad \text{if } |\mathbf{y}| = \beta\delta. \end{aligned} \tag{5.5.33}$$

Hence, if α is very close to α_c the right-hand sides of both of Eqs. (5.5.33) are negative. We use this to infer in a standard fashion that there is a function $\varepsilon_\delta > 0$ such that $\forall\,\alpha \in (a_c - \varepsilon_\delta, \alpha + \varepsilon_\delta)$, the set $\Gamma(\delta)$ is $S_t^{(\alpha)}$-invariant (where, as usual, the solution flow for our equation is denoted $S_t^{(\alpha)}$. In fact, let ε_δ be a monotonically decreasing function of $\delta \in (0, \delta_0]$ such that $\sigma(\alpha) < \frac{1}{4}|\overline{\gamma}|\delta^2$ for $\alpha \in (\alpha_c - \varepsilon_\delta, \alpha_c + \varepsilon_\delta)$. For such values of α_c the right-hand sides of both of Eqs. (5.5.33) are negative.

Then let $\mathbf{x} = (x_1, x_2, \mathbf{y}) \in \Gamma(\delta)$ and let $\overline{t} =$ (first time > 0 such that $S_t^{(\alpha)}(\mathbf{x}) \notin \Gamma(\delta)$ and note that either the first or the second of Eqs. (5.5.33) (according to which side of $\partial\Gamma(\delta)$ is crossed) implies that $S_t^{(\alpha)}(\mathbf{x}) \notin \Gamma(\delta)$ for some earlier time $t < \overline{t}$ against the definition of $\overline{t}$. So $\overline{t} = +\infty$ and $\Gamma(\delta)$ is invariant for $S_t^{(\alpha)}$, $\forall\, t \ge 0$, $\forall \in (\alpha_c - \varepsilon_\delta, \alpha_c + \varepsilon_\delta)$, $\forall\, \delta \le \delta_0$.

To prove vague attractivity, see Definition 5, and Observation (8), p.391, it is natural to try to choose $U = \Gamma(\delta_0)$. Therefore, ask the following question: given $\delta \le \delta_0$ and $\alpha \in (a_c - \varepsilon_\delta, \alpha_c + \varepsilon_\delta)$, can we find $t_\delta > 0$ such that $S_t^{(\alpha)}\Gamma(\delta_0) \subset \Gamma(\delta)$?

Let $\mathbf{x} \in \Gamma(\delta_0)/\Gamma(\delta)$ and suppose that for t in some interval $[0, T]$, $S_t^{(\alpha)}(\mathbf{x}) \in \Gamma(\delta_0)/\Gamma(\delta)$. If we define $\delta(t)^2 \stackrel{def}{=} \max(\varrho(t)^2, \mathbf{y}(t)^2/\beta^2)$, the point $S_t^{(\alpha)}(\mathbf{x})$ is in $\partial\Gamma(\delta(t))$ and Eq. (5.5.33), together with the assumption that $\forall\, t \in [0, T]$, $\delta \le \delta(t) \le \delta_0$, imply

$$\delta(t)^2 \le \delta(0)^2 + 2T\max(\frac{\overline{\gamma}}{4}\delta^2, -\frac{\nu}{2}\delta^2) \le \delta_0^2 - TM_\delta \tag{5.5.34}$$

with $M_\delta > 0$.[11] Hence, if $t_\delta = (\delta_0^2 - \delta^2)/M_\delta$, it follows that $T < t_\delta$. Hence, $S_t^{(\alpha)}\Gamma(\delta_0) \subset \Gamma(\delta)$, $\forall\, \alpha \in (\alpha_c - \varepsilon_\delta, \alpha_c + \varepsilon_\delta)$.

It is now clear that $\mathbf{x}_0 = 0$ is vaguely attractive. By Definition 5, and observation (8), p.391, one can take $U = \Gamma(\delta_0), \varrho_\delta, \forall\, \delta\lambda\delta_0, \varepsilon_\delta, t_\delta$ as above for all $\delta \le \delta_0$ and Eq. (5.5.7) holds for $\delta \le \delta_0$. If $\delta > \delta_0$, Eq. (5.5.7) follows from the supposed uniform boundedness of the trajectories,

So the proof of vague attractivity is complete as long as $\gamma_\alpha(\theta) \equiv \overline{\gamma}_\alpha$, $\forall\, \alpha \in (\alpha_c - a, \alpha_c + a)$. We must now remove this restriction.

This will be achieved by studying a coordinate change $(x_1, x_2, \mathbf{y}, \alpha) \to (\xi_1, \xi_2, \mathbf{y}', \alpha')$ with $\alpha' \equiv \alpha$, $\mathbf{y}' \equiv \mathbf{y}$, and

$$\begin{aligned}\xi_i =&x_i + \sum_{j,k=1}^{2} a_{ijk}(\alpha) x_j x_k, \qquad i = 1,2\\ x_i =&\xi_i - \sum_{j,k=1}^{2} a_{ijk}(\alpha)\xi_j \xi_k + H_i(\boldsymbol{\xi}, \alpha), \qquad i = 1,2,\end{aligned} \tag{5.5.35}$$

where H, a_{ijk} are C^∞ functions of their arguments and defined in the neighborhoods of $(\mathbf{0}, a_c)$ of the form $V \times I$, $V \subset \mathcal{R}^d$ open; furthermore H_i have a third order zero at $\boldsymbol{\xi} = \mathbf{0}$.

We must show that $a_{ijk}(\alpha)$ can be so chosen that the two equations

$$\begin{aligned}\dot{x}_1 =&\sigma(a)x_1 - \mu(\alpha)x_2 + \sum_{h,k=1}^{2} \overline{N}_{1hk\ell=1}(\alpha) x_h x_k x_\ell,\\ \dot{x}_2 =&\mu(a)x_1 + \sigma(\alpha)x_2 + \sum_{h,k=1}^{2} \overline{N}_{1hk\ell=1}(\alpha) x_h x_k x_\ell,\end{aligned} \tag{5.5.36}$$

[see Eqs. (5.5.28) and (5.5.29)] are changed into

---

[11] Here we use a lemma on integration theory: if $a, b > 0$ are two C^∞-functions bounded below by a positive constant $\sigma > 0$ and if $d(t) = \max(a(t), b(t))$ and $c(t) = \dot{a}(t)$ for $a(t) > b(t)$, $c(t) = \dot{b}(t)$ for $a(t) < b(t)$, and $c(t) = \frac{1}{2}(\dot{a}(t) + \dot{b}(t))$ if $a(t) = b(t)$, then $d(t) = d(0) + \int_0^t c(\tau)d\tau \le d(0) + t\sup_{0\le\tau\le t} c(\tau)$.

This can be proved by remarking that

$$\begin{aligned}d(t) =& \lim_{N\to\infty} (a(t)^N + b(t)^N)^{\frac{1}{N}} = d(0) + \lim \int_0^t \frac{d}{d\tau}(a(\tau)^N + b(\tau)^N)^{\frac{1}{N}} d\tau\\ =&d(0) + \lim_{N\to\infty} \int_0^t \frac{a(\tau)^{N-1}\dot{a}(\tau) + b(\tau)^{N-1}\dot{b}(\tau)}{(a(\tau)^N + b(\tau)^N)^{1-\frac{1}{N}}} d\tau\end{aligned}$$

and the function under the integration sign is uniformly bounded in N by the $\max_{0\le\tau\le t}(|\dot{a}(\tau)|, |\dot{b}(\tau)|)$ and it is pointwise convergent to $c(\tau)$. If $a(\tau) = b(\tau)$ has only a finite number of solutions τ, the possibility of taking the limit under the integral sign is easily proved. If $a(\tau) = b(\tau)$ has infinitely many roots one can find a simple approximation argument, recalling that a, b are bounded below by $\sigma > 0$, to infer $d(t) \le d(0) + t\sup_{0\le\tau\le t} c(\tau)$. Alternatively, one can apply the dominated convergence theorem of Lebesgue.

$$\begin{aligned}\dot{\xi}_1 =& \sigma(a)\xi_1 - \mu(\alpha)\xi_2 + \overline{\gamma}_\alpha \xi_1 \, (\xi_1^2 + \xi_2^2) + \text{4-th order terms} \\ \dot{\xi}_2 =& \mu(a)\xi_1 + \sigma(\alpha)\xi_2 + \overline{\gamma}_\alpha \xi_2 \, (\xi_1^2 + \xi_2^2) + \text{4-th order terms}\end{aligned} \tag{5.5.37}$$

In fact, such a change of variables would manifestly change Eq. (5.5.27) into an equation of the same normal form but with $\gamma_\alpha(\theta) \equiv \overline{\gamma}_\alpha$.

The existence of such a change of coordinates is easier to discuss after introducing the variables $z \stackrel{def}{=} x_1 + i\, x_2, \overline{z} = x_1 - i\, x_2$, $\lambda \stackrel{def}{=} \sigma + i\, \mu$, $\zeta \stackrel{def}{=} \xi_1 + i\, \xi_2$, $\overline{\zeta} = \xi_1 - i\, \xi_2$ and writing Eq. (5.5.36) as an equation for z, multiplying the second equation by i and adding it to the first (see [29]):

$$\dot{z} = \lambda(\alpha)z + a_3(\alpha)z^3 + a_2(\alpha)z^2\overline{z} + a_1(\alpha)z\overline{z}^2 + a_0(\alpha)\overline{z}^3, \tag{5.5.38}$$

where $a_0, \ldots, a_3$ are complex numbers that can be obtained from the N's by suitable linear combinations. Similarly Eq. (5.5.35) in complex form is:

$$\begin{aligned}\zeta =& z + A_3(\alpha)z^3 + A_2(\alpha)\zeta^2\overline{z} + A_1(\alpha)z\overline{z}^2 + A_0(\alpha)\overline{z}^3, \\ z =& \zeta - A_3(\alpha)\zeta^3 - A_2(\alpha)\zeta^2\overline{z} - A_1(\alpha)\zeta\overline{\zeta}^2 - A_0(\alpha)\overline{\zeta}^3 + H(\zeta, \overline{\zeta}, \alpha)\end{aligned} \tag{5.5.39}$$

with H having a zero of fourth order in $|\zeta|$ as $\zeta \to 0$, $\forall\, \alpha \in (\alpha_c - a, \alpha_c + a)$. Note that Eqs. (5.5.38) and (5.5.39) also imply an expression for $\gamma_\alpha(\theta)$: if $z = \varrho e^{i\theta}$, then

$$\begin{aligned}\gamma_\alpha(\theta) =& \varrho^{-4} \mathcal{R}\mathrm{e}\, (a_3(\alpha)\overline{z}z^3 + a_2(\alpha)z^2\overline{z}^2 + a_1(\alpha)z\overline{z}^3 + a_0(\alpha)\overline{z}^4) \\ =& \mathcal{R}\mathrm{e}\, (a_3(\alpha)e^{2i\theta} + a_2(\alpha) + a_1(\alpha)e^{-2i\theta} + a_0(\alpha)\overline{e}^{-4i\theta})\end{aligned} \tag{5.5.40}$$

which follows after some algebra, starting with the observation that $(x_1N_1 + x_2N_2) = \mathcal{R}\mathrm{e}\,(\overline{z}N)$, if $N \stackrel{def}{=} N_1 + iN_2$ denotes the complex combination of the nonlinear terms in the right-hand side of Eq. (5.5.36). Hence,

$$\overline{\gamma}_\alpha = \frac{1}{2\pi}\int_0^{2\pi} \gamma_\alpha(\theta)d\theta = \mathcal{R}\mathrm{e}\, a_2(\alpha)$$

So the goal is to determine A_3, A_2, A_1, A_0 in Eq. (5.5.39) so that (5.5.38) in the ζ variables has a third-order term of the form $a_2(\alpha)\zeta^2\overline{\zeta}$. A calculation shows that the equation for ζ is

$$\begin{aligned}\dot{\zeta} =& \dot{z} + 3A_3 z^3\dot{z} + 2A_2 z\overline{z}\dot{z} + A_2 z^2\dot{\overline{z}} + A_1\dot{z}\overline{z}^2 + 2A_1\overline{z}\dot{\overline{z}} + 3A_0\overline{z}^2\dot{\overline{\zeta}} \\ =& \lambda\zeta + \zeta^3(a_3 + 2\lambda A_3) + \zeta^2\overline{\zeta}(a_2 + (\lambda + \overline{\lambda})A_2) + \zeta\overline{\zeta}^2(a_1 + 2\overline{\lambda}A_1) \\ &+ \overline{\zeta}^3(a_0 - (\lambda - 3\overline{\lambda})A_0) + \text{4-th order terms}\end{aligned} \tag{5.5.41}$$

hence, we take

$$A_3 = -\frac{a_2}{2\lambda},\ a_2 = 0,\ A_1 = -\frac{a_1}{2\overline{\lambda}},\ A_0 = \frac{a_0}{\lambda - 3\overline{\lambda}} \tag{5.5.42}$$

and, near a_c the equation becomes

$$\dot{\zeta} = \lambda(\alpha)\zeta + a_2(\alpha)\overline{\zeta}\zeta^2 + \text{4-th order terms} \tag{5.5.43}$$

whose $\gamma_\alpha(\theta)$ function is $\gamma_\alpha(\theta) \equiv \gamma_\alpha = \mathcal{R}\mathrm{e}\, a_2(\alpha)$.

The proof that if $\overline{\gamma} > 0$ the origin is not vaguely attractive is left as a problem for the reader. mbe

It may be interesting to state explicitly some elementary invariance criteria for sets, which have been implicitly proved in the course of the above proof of Proposition 8.

9 Proposition. *(i) Let $U \subset \mathcal{R}^d$ be an open set with regular boundary ∂U. Then U is invariant for $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{f} \in C^\infty(\mathcal{R})$, if*

$$\mathbf{f}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) < 0, \qquad \forall\ \mathbf{x} \in \partial U, \tag{5.5.44}$$

were $\mathbf{n}(\mathbf{x})$ is the outer normal to ∂U in $\mathbf{x}$.
(ii) Let $V \in C^1(\mathcal{R}^d)$ and let $U(\mu) = \{\mathbf{x}\,|\,\mathbf{x} \in \mathcal{R}^d,\ V(\mathbf{x}) < \mu\}$. If $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{f} \in C^\infty(\mathcal{R}^d)$, is a differential equation and

$$\frac{\partial V}{\partial \mathbf{x}}(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) < 0, \qquad \forall\ \mathbf{x} \in \partial U(\mu), \tag{5.5.45}$$

then the set $U(\mu)$ is invariant. Furthermore, if

$$\sup_{V(\mathbf{x}) \in (\mu_1,\mu_2)} \frac{\partial V}{\partial \mathbf{x}}(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) = -C < 0 \tag{5.5.46}$$

and if $\mu' < \mu''$, $[\mu', \mu''] \subset (\mu_1, \mu_2)$, then

$$S_t U(\mu'') \subset U(\mu'), \qquad \forall\ t > \frac{\mu'' - \mu'}{C}. \tag{5.5.47}$$

Observations
(1)This proposition can be extended to the case when V is "piecewise C^∞ by replacing $\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}$ with the set of the convex linear combinations of its extreme values (i.e., by a suitable bundle of vectors "pointing out of $U(V(\mathbf{x}))$" in $\mathbf{x}$). This is useful because sometimes V may have a square or a cylinder as its level surface, as was the case for $\Gamma(\delta)$ after Eq. (5.5.33), where $V(\mathbf{x}) = \max(x_1^2 + x_2^2)$.

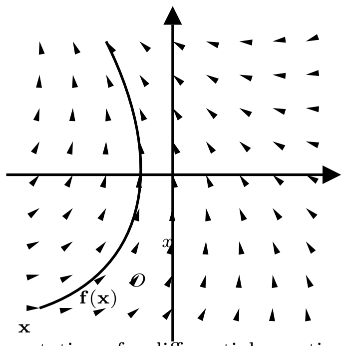


Figure 5.2 Geometric interpretation of a differential equation as a vector field.

(2) The geometric interpretation of a differential equation is the following: at every $\mathbf{w} \in \mathcal{R}^d$, draw a vector $\mathbf{f}(\mathbf{w})$, i.e., think of $\mathbf{f}$ as a "vector field" over $\mathcal{R}^d$. A solution to $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is associated with a curve in $\mathcal{R}^d$ which at every point is tangent to the vector field at the same point. This curve is run at a speed which at every point is equal to the modulus of the field vector at that point and has the same direction, see Fig. 5.2.
(3) The first statement of Proposition 9 is illustrated in Fig. 5.3.

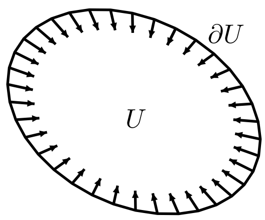


Figure 5.3: A vector field at the boundary of an invariant set U which implies its invariance.

Observation (1) is illustrated in Fig. 5.4:

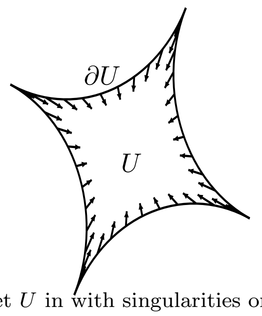


Figure 5.4: As in Fig.5.3 for a set U in with singularities on the boundary.

In connection with the above remarks, it is useful to see some pictures of a vaguely attractive point for an equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$. The "loss of stability" of a fixed point for a differential equation depending on a parameter α consequence of the crossing of the imaginary axis of some eigenvalues of the stability matrix as α varies is called a *bifurcation*: hence the above vague attractivity analysis deals with examples of vagueley attractive bifurcations.

In Figs. 5.5, we draw the vector field for α slightly larger than α_c, in a typical case of a vaguely attractive bifurcation of the origin.

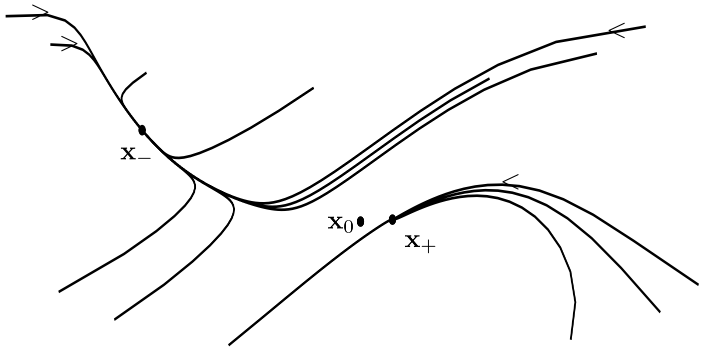


Figure 5.5 A vector field following a vaguely attractive bifurcation in one real direction: two attractive fixed points appear $\mathbf{x}_\pm$ and the bifurcating point remains as a repulsive fixed point ($\mathbf{x}_0$).

In Fig. 5.6 a vector field following a bifurcation of the rigid with vaguely attractive loss of stability with two imaginary eigenvalues

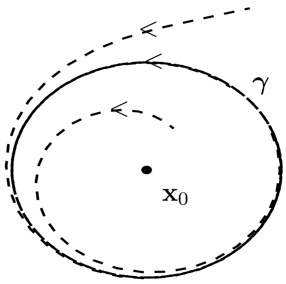


Figure 5.6 A vector field following a vaguely attractive bifurcation in one complex direction: a periodic orbit γ appears (solid line) and trajectories starting close to $\boldsymbol{\gamma}$ inside or outside it spiral towards it (dashed lines) and the bifurcating point remains as a repulsive fixed point ($\mathbf{x}_0$).

The reader should try to understand such pictures by trying to draw them on the basis of the above information and comments on how a vector field should look near a vaguely attractive point.

Figures 5.5 and 5.6 allow one to see immediately that in the vicinity of a vaguely attractive fixed point, there should usually appear two fixed points, Fig. 5.5, or a periodic orbit, Fig. 5.6, depending on whether the stability loss takes place, as α passes through α_c in one real direction or in two complex-conjugate directions.

In fact, this is the essential content of the Hadamard-Perron theorem and of the Hopf theorem which we will discuss in the upcoming sections.

We conclude this section by returning to the problem that we have been using to motivate the analysis of this section: the stability of the stationary solution $\widehat{\boldsymbol{\omega}}$ of Eq. (5.1.19).

10 Proposition. *Consider Eq. (5.1.19). The stationary solution* $\widehat{\boldsymbol{\omega}}$ *is vaguely attractive near* $a_c = \lambda_1 + \lambda_2''' \,\widehat{\omega}_3^2$.

Observation. The following proof shows that one should not blindly begin to compute mechanically the vague-attractivity constant $\overline{\gamma}_{\alpha_c}$. The reader will note the use of several "tricks" which are not worth being organized in a sequence of propositions refining the criterion of Proposition 8, but which, nevertheless, make the computation reasonably short. The reader should use these tricks in the exercises at the end of this section.

PROOF. We apply Proposition 8, p.396. Changing variables to bring the fixed point to the origin, i.e., $(\omega_1, \omega_2, \omega_3) \longleftrightarrow (\varrho_1, \omega_2, r)$, $r = \omega_3 - \widehat{\omega}_3$, Eq. (5.1.19) becomes, setting $\widetilde{\lambda}_2 \stackrel{def}{=} (\lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2'''r^2)$,

$$\begin{aligned}
\dot{\omega}_1 =& (\alpha - \alpha_c)\omega_1 - \widehat{\omega}_3\omega_3 - 2\lambda_2'''\widehat{\omega}_3\, r\, \omega_1 - \omega_2 r - \widetilde{\lambda}_2\omega_1, \\
\dot{\omega}_2 =& \widehat{\omega}_3\omega_1 + (\alpha - \alpha_c)\omega_2 - 2\lambda_2'''\widehat{\omega}_3\, r\, \omega_2 + \omega_1 r - \widetilde{\lambda}_2\omega_2, \\
\dot{r} =& -(\lambda_1 + 3\lambda_2'''\widehat{\omega}_3^2)\, r\, - \widehat{\omega}_3\, \widetilde{\lambda}_2 - \widetilde{\lambda}_2 r.
\end{aligned} \tag{5.5.48}$$

It is convenient to condense Eq. (5.5.48) by introducing

$$\begin{aligned}
z \stackrel{def}{=}& \omega_1 + i\omega_2, \;\; \lambda \stackrel{def}{=} (\alpha - \alpha_c) + i\widehat{\omega}_3, \\
\widetilde{\lambda} \stackrel{def}{=}& -(\lambda_1 + 3\lambda_2'''\widehat{\omega}_3^2), \;\; E \stackrel{def}{=} 2\lambda_2'''\widehat{\omega}_3 - i, \\
Q(r, z, \overline{z}) \stackrel{def}{=}& \lambda_2'(\mathcal{R}\mathrm{e}\, z)^2 + \lambda_2''(\mathcal{I}\mathrm{m}\, z)^2 + \lambda_2'''r^2 \\
P(r, z, \overline{z}) \stackrel{def}{=}& (\lambda_2'(\mathcal{R}\mathrm{e}\, z)^2 + \lambda_2''(\mathcal{I}\mathrm{m}\, z)^2 + 3\lambda_2'''r^2)\widehat{\omega}_3
\end{aligned} \tag{5.5.49}$$

Then, multiplying the second of Eqs. (5.5.48) by i and adding it to the first, we find that Eq. (5.5.48) becomes

$$\dot{z} = (\lambda - Er - Q)z, \;\; \dot{r} = \widetilde{\lambda}r - P - r\,Q. \tag{5.5.50}$$

To put the above equation in normal form with respect to $\lambda, \overline{\lambda}$ change variables (see Proposition 7) as:

$$\zeta = z - A\,z\,r, \qquad z = \frac{\zeta}{1 - A\,r} \tag{5.5.51}$$

Then the first of Eqs. (5.5.49) becomes

$$\begin{aligned}
\dot{\zeta} =& z\,(\lambda - Er - Q) - A\,r\,z\,(\lambda - Er - Q) - A\,z\,(\widetilde{\lambda}r - P - r\,Q) \\
=& \zeta\, \frac{(\lambda - Er - Q)(1 - A\,r) - A\,(\widetilde{\lambda}r - P - r\,Q)}{1 - A\,r} \stackrel{def}{=} \zeta\, F(\zeta),
\end{aligned} \tag{5.5.52}$$

whose linear and quadratic terms are $\lambda\zeta$ and $-(\lambda Ar + Er + \widetilde{\lambda}Ar - \lambda Ar)\zeta \equiv -(Er + \widetilde{\lambda}Ar)\zeta$, respectively . So we choose $A = -E/\widetilde{\lambda}$ and Eq. (5.5.48) acquires

normal form in the (ζ, r) variables with respect to the eigenvalues $\lambda, \overline{\lambda}$. From Eq. (5.5.52), it is then easy to compute $\gamma_\alpha(\theta)$: if $\zeta = \varrho e^{i\theta}$,

$$\gamma_\alpha(\theta) = \lim_{\varrho\to 0} \Big[\mathcal{R}\mathrm{e}\, \frac{|\zeta|^2 (F(\zeta) - \lambda)}{\varrho^4}\Big]_{r-0} = \lim_{\varrho\to 0} \mathcal{R}\mathrm{e}\,(\lambda - Q_0 + AP_0 - \lambda)\varrho^{-4} \tag{5.5.53}$$

if Q_0, P_0 are Q, P with $r = 0$. Therefore,

$$\begin{aligned}\gamma_\alpha(\theta) &= -\,(\lambda_2' \cos^2\theta + \lambda_2'' \sin^2\theta) + \frac{2\lambda_2''' \widehat{\omega}_3^2 (\lambda_2' \cos^2\theta + \lambda_2'' \sin^2\tau)}{\lambda_1 + 3\lambda_2''' \widehat{\omega}_3^2} \\ &= -\,(\lambda_2' \cos^2\theta + \lambda_2'' \sin^2\theta) \frac{\lambda_1 + \lambda_2''' \widehat{\omega}_3^2}{\lambda_1 + 3\lambda_2''' \widehat{\omega}_3^2}\end{aligned} \tag{5.5.54}$$

which yields

$$\overline{\gamma}_\alpha = -\frac{\lambda_2' + \lambda_2''}{2} \frac{\lambda_1 + \lambda_2''' \widehat{\omega}_3^2}{\lambda_1 + 3\lambda_2''' \widehat{\omega}_3^2} \tag{5.5.55}$$

and Proposition 10 follows from Proposition 8. mbe

5.5.1 Exercises

1. Study the vague attractivity of the fixed point $x = 0$ of $\dot{x} = \alpha x + f(x)$, where $f \in C^\infty(\mathcal{R}), f(0) = 0, f'(0) = 0$. Show that the origin cannot be vaguely attractive unless $f''(0) = 0$.

2. Show, by producing some examples, that if the number $\overline{\gamma}_{\alpha_c}$, of Proposition 8 vanishes, then the fixed point may or may not be vaguely attractive. (*Hint:* Find examples other than $i\dot{z} = (\alpha + i\mu)z \pm z^3 z^2, \alpha \in \mathcal{R}, \mu \in \mathcal{R}, \alpha_c = 0, z \in \mathcal{C}$.)

3. Suppose that the origin is vaguely attractive for $\dot{x} = x\, f(x^2, \alpha)$ near α_c. Show that the equation

$$\dot{x} = -\mu y + x\, f(x^2 + y^2, \alpha), \qquad \dot{y} = \mu x + y\, f(x^2 + y^2, \alpha),$$

also has the origin as a vague attractor near α_c, $\forall\, \mu \in \mathcal{R}$.

4. Let $a_1 \in C^\infty(\mathcal{R})$. Show that the origin is a vague attractor near α_c for $\dot{x} = -x\,(x^2 - a_1(\alpha))$ if $a_1(\alpha_c) = 0$.

5. Given $a_1 \in C^\infty(\mathcal{R})$ show that the origin is vaguely attractive for $\dot{x} = -x\,(x^2 - a_1(\alpha))(x^2 - a_2(\alpha))$ near α_c if $a_1(\alpha_c) = a_2(\alpha_c) = 0$. (*Hint:* use Observation (2) to Definition 5, p.391.)

6. Compute the vague-attractivity indicator $\overline{\gamma} = \overline{\gamma}_{\alpha_c}$ for the origin, see Proposition 8, in the equation

$$\dot{x} = -\mu y - x\,(x^2 + y^2 - a_1(\alpha)), \qquad \dot{y} = \mu x - y\,(x^2 + y^2 - a_1(\alpha)),$$

assuming $\mu \in \mathcal{R}, a_1(\alpha_c) = 0, a_1 \in C^\infty(\mathcal{R})$.

7. Same as Problem 6 for

$$\dot{x} = -\mu y - x\,(x^2 + y^2 - a_1(\alpha))(x^2 + y^2 - a_2(a)),$$
$$\dot{y} = \mu x - y\,(x^2 + y^2 - a_1(\alpha))(x^2 + y^2 - a_2(a)),$$

assuming $\mu \in \mathcal{R}, a_1, a_2 \in C^\infty(\mathcal{R}), a_1(_c) = a_2(\alpha_c) = 0$, study the vague attractivity. (From [35]).

8. Let $z = x_1 + ix_2$, $\lambda = \alpha + i\mu, , \alpha, \mu \in \mathcal{R}$, and consider the differential equation $\dot{z} = \lambda z - \alpha a z\overline{z} - z^2\overline{z}$. Apply Proposition 8 to find the vague-attractivity indicator for the origin near $\alpha_c = 0$. (*Answer:* $\overline{\gamma} = -1$). What can be said about the vague attractivity of the origin when $\mu = 0$? (*Warning:* Note that the equation does not have normal form.)

9. Under the assumptions of the first sentence of Proposition 8, only suppose that the equation $\dot{\mathbf{x}}x = \mathbf{f}(\mathbf{x}, \alpha)$ has the form

$$\dot{x}_1 = \sigma(\alpha)x_1 - \mu(\alpha)x_2 + S_1(x_1, x_2, \mathbf{y}, \alpha) + N_1(x_1, x_2, \mathbf{y}, \alpha)$$
$$\dot{x}_2 = \mu(\alpha)x_1 - \sigma(\alpha)x_2 + S_2(x_1, x_2, \mathbf{y}, \alpha) + N_2(x_1, x_2, \mathbf{y}, \alpha)$$
$$\dot{\mathbf{y}} = \widetilde{L}(\alpha)\mathbf{y} + \mathbf{F}(x_1, x_2, \mathbf{y}, \alpha),$$

where $\lambda_1(\alpha) = \overline{\lambda}_2(\alpha) = \sigma(\alpha) + i\,(\alpha)$, $\mathbf{F}$ has second-order zero at the origin $x_1 = x_2 = 0, \mathbf{y} = \mathbf{0}$, for all $\alpha \in I$, N_1 and N_2 have a third order zero at the origin for all $\alpha \in I$, and S_1, S_2 are homogeneous second-order polynomials in $x_1, x_2, \mathbf{y}$. All the functions are supposed to be of class C^∞ in their arguments $(x_1, x_2, \mathbf{y}, \alpha) \in \mathcal{R}^d \times I$.
Suppose, furthermore, $\mathbf{F}(x_1, x_2, \mathbf{0}, \alpha) \equiv \mathbf{0}$ and also $S_1(x_1, x_2, \mathbf{0}, \alpha_c) \equiv S_2(x_1, x_2, \mathbf{0}, \alpha_c) = 0$ and define $\gamma_\alpha(\theta)$ by Eq. (5.5.24) with the present meaning of the symbols. Show that the origin is vaguely attractive near a point $\alpha_c \in I$, where $\sigma(\alpha_c) = 0, \mu(\alpha_c) \neq 0$, if $\overline{\gamma} < 0$, in spite of the presence of the terms S_1, S_2. (*Hint:* Show that the above equation can be put into normal form with respect to $\lambda_1, \overline{\lambda}_1$ with a change of variables like

$$\xi_j = x_j + \sum_{h=1}^{2}\sum_{k=1}^{d-2} A_{jkh}x_h y_k + \sum_{h,k=1}^{d-2} B_{jkh}y_h y_k, \qquad k, j = 1, 2,$$

and this change of variables does not affect the value of $\gamma_\alpha(\theta)$ because it changes the third-order terms by a quantity vanishing as $\mathbf{y} \to \mathbf{0}$.) This extends Problem 8.

10. Prove that the same conclusions of Proposition 8 hold, replacing the assumption that N_1 and N_2 are of third order with the assumption that $x_1N_1 + x_2N_2$ has a fourth-order zero at $x_1 = x_2 = 0, \mathbf{y} = \mathbf{0}$, for all $\alpha \in I$. (*Hint:* Simply go through the proof of Proposition 8.)

11. Same as Problem 9, replacing the assumption that N_1 and N_2 are of third order with the assumption that N_1 and N_2 is of fourth order.

12. Show that the origin is not a vaguely attractive point near $a_c = 0$ for all the values of $E \in \mathcal{C}$ in the equation in $\mathcal{R}^3$:

$$\dot{z} = \lambda z + Ezr - z\overline{z}^2, \quad \dot{r} = -r + z\overline{z},$$

where $z = x + iy, \lambda = \alpha + i\mu, \mu \neq 0,\ x, y, r \in \mathcal{R}^3$.

13. Put into normal form the equation

$$\dot{\omega}_1 = -\omega_1 - \omega_2 + \omega_1\omega_2, \qquad \dot{\omega}_2 = \omega_1 - \omega_2 + \omega_1^2.$$

(*Hint:* Introduce $z = \omega_1 + i\omega_2$ and change variables as $\zeta = z + A_2z^2 + A_1z\overline{z} + A_0\overline{z}^2$, etc.)

14. Analyze the vague attractivity of the stationary solution $\omega_1 = \omega_2 = 0, \omega_3 = 5\alpha$ of the equation

$$\dot{\omega}_1 = -\omega_1 - \omega_2\omega_3, \quad \dot{\omega}_2 = -\frac{1}{3}\omega_2 + \omega_1\omega_3, \quad \omega_3 = -\frac{1}{5}\omega_3 + \alpha.$$

15. Consider the equation ("Lorenz equation")

$$\dot{x} = \sigma(-x + y), \quad \dot{y} = -\sigma x - y - xz, \quad \dot{z} = -bz + xy - \alpha$$

and study the vague attractivity of its fixed points for $b = \sigma = 1$. (*Hint:* For the analysis of the fixed point $z = -\frac{\alpha}{b}$ at $\alpha_c = (1+\sigma)b$ try to use Eqs. (5.5.18) and (5.5.11), with $\alpha = \alpha_c, j = 1$, to put the first equation (in the appropriate variables) into the normal form of Eq. (5.5.26); the result will be $\overline{\gamma} = -\sigma/b$. *Warning*: The analysis of the other fixed points is very cumbersome.)

16. Same as Problem 15 for $b = \frac{8}{3}$, $\sigma = 10$.

17. Suppose that the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ has a stationary solution $\mathbf{x}(\alpha)$ for $le\alpha_c$, depending continuously on α. Let $L(\alpha), \lambda_1(\alpha), \ldots, \lambda_n(\alpha)$ be the stability matrix and its eigenvalues. Assume that, for $\alpha < \alpha_c$, it is $\mathcal{R}\mathrm{e}\,\lambda_j(\alpha) < 0$ and, for $\alpha = a_c$ $Re\lambda_{j_0}(\alpha_c) = 0$ for some j_0.

Show that if no eigenvalue actually vanishes at $\alpha = \alpha_c$, (i.e., $\mathcal{I}\mathrm{m}\,\lambda_j(\alpha_c) \neq 0$, $j = 1, \ldots, n$), then the solution $\mathbf{x}(\alpha)$ can be continuously continued to $\alpha \geq \alpha_c$, (i.e., there is a continuous function a $\to \mathbf{x}(a)$ defined in the vicinity of α_c, and $\mathbf{f}(\mathbf{x}(a)) \equiv 0$).

Show also that if there is an eigenvalue vanishing at $a = a_c$ the solution $\mathbf{x}(\alpha)$ will not admit, in general, a continuation for $\alpha > \alpha_c$. (*Hint:* Just use the implicit functions theorem for the equation $\mathbf{f}(\mathbf{x}, \alpha) = \mathbf{0}$ near$(x(\alpha_c), \alpha_c)$; then consider the example $f(x) = ax + x^2 + \alpha$, $x(a) = (-\alpha - (-4\alpha + \alpha^2)^{\frac{1}{2}})/2$, $a_c = 0, L(a) \equiv \lambda(\alpha) \simeq -\sqrt{-4\alpha}$.)

5.6 Vague-Attractivity Properties. The Attractive Manifold

> Every five years or so, if not more often, someone discovers
> the theorem of Hadamard and Perron, proving it by
> Hadamard's method or by Perron's. (Anosov)

The solution $\widehat{\boldsymbol{\omega}}$ of Eq. (5.1.19), thought of as a family of differential equations parameterized by a parameter α, is, as shown in Proposition 10, p.405, §5.5, vaguely attractive near $\alpha_c = \lambda_1 + \lambda_2''' \widehat{\omega}_3^2$.

Therefore, the motion $t \to S_t^{(\alpha)}(\boldsymbol{\omega}), t \geq 0$, with initial datum close to $\widehat{\boldsymbol{\omega}}$ continues to remain quite close to $\widehat{\boldsymbol{\omega}}$ if α is near α_c in spite of the instability of $\widehat{\boldsymbol{\omega}}$ for $\alpha > \alpha_c$. We shall see that $\widehat{\boldsymbol{\omega}}$ is not only unstable, but it also cannot be an attractor. Hence, the motions which develop from a datum in the vicinity of $\widehat{\boldsymbol{\omega}}$, although remaining there, cannot generally have an asymptotic behavior, as $t \to +\infty$, simply given by $S_t(\boldsymbol{\omega}) \to \widehat{\boldsymbol{\omega}}$.

In the linear approximation, when the right-hand side of Eq. (5.1.19) is replaced by the function $\boldsymbol{\omega} \to L_\alpha(\boldsymbol{\omega} - \widehat{\boldsymbol{\omega}})$, where L_α is the stability matrix (5.5.1) of Eq. (5.1.19) at $\widehat{\boldsymbol{\omega}}$, the motion is very simple and $\omega_3 \to \widehat{\omega}_3$ exponentially fast ($\simeq e^{-\lambda_1 t}$), while $\omega_1^2 + \omega_2^2$ grows exponentially (roughly as $e^{(\alpha - \alpha_c)t}$).

The linear approximation is certainly incorrect as soon as $\omega_1^2+\omega_2^2$ becomes large, just because $\widehat{\boldsymbol{\omega}}$ is vaguely attractive [see Observation (2) to Definition 5, p.391]. However, one can hope that even in the essentially nonlinear motion governed by Eq. (5.1.19), some memory remains of the fact that $\widehat{\boldsymbol{\omega}}$ lost stability" only in two directions, i.e., only for what concerns the components of the motion in the plane $\omega_3 = \widehat{\omega}_3$ generated by the eigenvectors $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}$ of the stability matrix of $\widehat{\boldsymbol{\omega}}$. We can then think that the motion following a given initial datum $\boldsymbol{\omega}$ close to $\widehat{\boldsymbol{\omega}}$ develops essentially on a two-dimensional surface, i.e., that the third a component $\omega_3(t)$ asymptotically tends to become a function $\varphi(\omega_1(t), \omega_2(t))$ of the first two.

More generally, we can imagine to find ourselves in the following situation, to which the upcoming Proposition 11 will refer.

Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ be a $\mathcal{R}^d$-valued function in $C^\infty(\mathcal{R}^d \times R)$ such that the differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha), \tag{5.6.1}$$

parameterized by α, have uniformly bounded trajectories as α varies in $I = (\alpha_c - a, \alpha_c + a)$, $a \in (0, 1)$, and have the point $\mathbf{x} = \mathbf{x}_0$ as a stationary solution, $\forall\, \alpha \in I$,

$$\mathbf{f}(\mathbf{x}_0, \alpha) = \mathbf{0}, \qquad \forall\ \alpha \in I. \tag{5.6.2}$$

Let L_α be the stability matrix in $\mathbf{x}_0$ and suppose that for $\alpha < \alpha_c$, all its eigenvalues $\lambda_1(\alpha), \dots, \lambda_d(\alpha)$ have a negative real part, while for $\alpha \in I = (\alpha_c - a, \alpha_c + a)$, only $d - r$ eigenvalues have real parts less or equal to $-\nu_0 < 0$, the others having real parts larger or equal than $-\nu_0' > -\nu_0$ and vanishing for $\alpha = \alpha_c$ (i.e., $\mathcal{R}\mathrm{e}\, \lambda_j(\alpha) = 0$ for $j = 1, \dots, r$).

For simplicity, also suppose that the eigenvalues of L_α are pairwise distinct: so we can choose the eigenvalues $\lambda_1(\alpha), \dots, \lambda_d(\alpha)$ and the corresponding eigenvectors $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(d)}$ so that they are $C^\infty(I)$ functions of α and so that every complex eigenvector appears together with a complex conjugate eigenvector. We suppose that the eigenvectors and eigenvalues have been so chosen and enumerated.

Under the above assumptions, we may assume without further loss of generality that for $\alpha \in I$, the equation takes the form of Eq. (5.5.10) and that the first r equations describe the evolution of the coordinates relative to the real plane generated by the "unstable" directions $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(r)}$ corresponding to the eigenvalues with large real part ($> -\nu_0'$). Were this not true, we could change the coordinates near $\mathbf{x}_0$ (see Proposition 8, p.396) to make this true. Finally, suppose $\mathbf{x}_0$ vaguely attractive for Eq. (5.6.1) near α_c and denote

$$U,\, \Gamma(\delta) \tag{5.6.3}$$

a system of neighborhoods associated to $\mathbf{x}_0$ for $\alpha \in I$, whose existence is guaranteed by Definition 5, p.390, of vague attractivity.

From the discussion of the preceding section, it appears that the just described situation can be realized for the Eq. (5.1.19), see Proposition 10, p.405, §5.5, thought of as a family of differential equations parameterized by α: thus this provides a concrete example to which the following theory can be applied.

We now formulate a proposition giving a positive answer to the conjecture hinted at above, that given $\alpha > \alpha_c$ close enough to α_c any motion of Eq. (5.6.1) starting close enough to $\mathbf{x}_0$ remains close to $\mathbf{x}_0$ (since $\mathbf{x}_0$ is vaguely attractive) and, furthermore, it can be thought of as developing asymptotically, for $t \to +\infty$, on an invariant surface σ_α.

Such a surface will have dimension r and it will be tangent to the "instability's hyperplane," $x_{r+1} = \ldots = x_d = 0$; furthermore, it will be an attractor for the motions starting in U and its attraction strength will be exponential and roughly measured, as in the linear case, by the parameter

$$-\nu = \max_{i=r+1,\ldots,d} \mathcal{R}\mathrm{e}\, \lambda_j$$

The surface σ_α will generally be non unique since, as we shall see, it may contain other smaller attractors: if Λ_α is a minimal attractor in σ_α which has U as its attraction basin, then, clearly, every invariant hypersurface $\sigma' \subset U$ containing Λ_α is an attractor for U; see the exercises at the end of this section.

Finally, the surface σ_α will be described inside the neighborhood $\Gamma(\mathbf{x}_0, \delta) = \{\text{cube centered in } \mathbf{x}_0 \text{ and side } 2\delta\}$ by $d-r$ functions on $\mathcal{R}^r \times \mathcal{R}$ of *preassigned* regularity $C^{(k)}$, $(k = 0, 1, \ldots)$, via the equations

$$x_{r+1} = \varphi^{(r+1)}(x_1, \ldots, x_r, \alpha), \ldots, x_d = \varphi^{(d)}(x_1, \ldots, x_r, \alpha), \ldots, \tag{5.6.4}$$

(where we suppose $\mathbf{x}_0 = \mathbf{0}$) provided δ is small and α is close to α_c. For α close to α_c this surface will be almost flat: if $k \geq 2$, this means that the first derivatives of the functions in Eq. (5.6.4) vanish for $(x_1, \ldots, x_r) = (x_{o1}, \ldots, x_{0r})$.

The interest in the above considerations is that it will become possible to analyze the asymptotic behavior of some properties of the motions originating near a vaguely attractive point $\mathbf{x}_0$, as $t \to +\infty$, reducing the d equations of Eq. (5.6.1) to the r equations, labeled by $j = 1, 2, \ldots, r$,

$$\dot{x}_j = f^{(j)}(x_1, \ldots, x_r, \varphi^{(r+1)}(x_1, \ldots, x_r, \alpha), \ldots, x_d = \varphi^{(d)}(x_1, \ldots, x_r, \alpha), \alpha). \tag{5.6.5}$$

When d is large and r is small, this may be a very important simplification.

When $r = 1$ or $r = 2$, this will say that the motion near the vaguely attractive point is a "one-dimensional" or "two-dimensional" problem. Figures 5.5-5.6 already suggest that in such cases it will be possible to obtain deeper insights into the theory of the asymptotic behavior of the solutions of the equations starting with initial data close to $\mathbf{x}_0$. They even suggest the results of such a theory (see Figs. 5.5 and 5.6 and §5.7).

In the case of Eq. (5.1.19), it is $r = 2$ and, therefore, the three equations (5.1.19) can be reduced, for $\alpha - \alpha_c$ small and for the purposes of the analysis of some asymptotic properties, to the first two equations with ω_3 replaced by

$$\omega_3 = \varphi(\omega_1, \omega_2, \alpha) + \widehat{\omega}_3 \tag{5.6.6}$$

where φ is a suitable $C^{(k)}$ function (with a preassigned k) having a second-order zero in $\omega_1, \omega_2 = 0$, i.e., such that there exist $\psi_1, \psi_2, \psi_3 \in C^{(k-2)}$ and

$$\varphi(\omega_1, \omega_2, \alpha) = \omega_1^2 \psi_1(\omega_1, \omega_2, \alpha) + \omega_2^2 \psi_2(\omega_1, \omega_2, \alpha) + \omega_1 \omega_3 \psi_3(\omega_1, \omega_2, \alpha) \tag{5.6.7}$$

expressing the tangency of the surfaces σ_α to the instability plane in $\widehat{\boldsymbol{\omega}}$, ($\omega_3 = \widehat{\omega}_3$ in this case) provided $k \geq 2$. For $k < 2$ the near flatness can be expressed by Eq.(5.6.11) below (implying Eq.(5.6.7) for $k \geq 2$).

A simple consequence of this, as we shall see in §5.7 (see footnote 15 on p.431), will be that for α close to α_c, $\alpha > \alpha_c$, there is a periodic orbit which is a normal and minimal attractor lying on σ_α with attraction basin $U/\mathcal{C}(\widehat{\boldsymbol{\omega}})$, where $\mathcal{C}(\widehat{\boldsymbol{\omega}})$ is a one-dimensional curve of points $\boldsymbol{\omega}$ through $\widehat{\boldsymbol{\omega}}$, whose asymptotic behavior is $S_t^{(\alpha)}(\boldsymbol{\omega}) \xrightarrow[t \to +\infty]{} \widehat{\boldsymbol{\omega}}$. Hence, it will be possible to draw a rather complete picture of the motion near $\widehat{\boldsymbol{\omega}}$.

A precise statement about the above matters is as follows.

11 Proposition. *Under the assumptions described in the above text between Eqs. (5.6.1) and (5.6.3), consider the symbols introduced there and let, for notational simplicity,* $\mathbf{x}_0 = \mathbf{0}$.

Given $k \geq 0, C > 0$, *there exist positive constants* $a_+, \delta, \delta_0, \nu \in (0,1)$ *with* $\delta_0 < \delta, a_+ < a$ *and* $d - r$ *functions of class* $C^{(k)}$, *denoted* $\varphi^{(k+1)} \ldots, \varphi^{(d)}$, *of the* $r+1$ *variables* $x_1, \ldots, x_r, \alpha$ *defined for*

$$|x_i| < \frac{\delta}{2}, \quad i = 1, \ldots, r, \quad \alpha \in I \equiv (\alpha_c - a_+, \alpha_c + a_+) \tag{5.6.8}$$

such that the surfaces $\sigma_\alpha \subset \mathcal{R}^d$ *described by Eqs. (5.6.4) have for all* $\alpha \in I$ *the properties:*
(i) "local invariance":

$$S_t^{(\alpha)}(\sigma_\alpha \cap \Gamma(\delta_0)) \subset \sigma_\alpha, \qquad \forall\, t \geq 0; \tag{5.6.9}$$

(ii) "local attractivity": there exist $C' > 0$ *such that for all* $\mathbf{w} \in U$ *it is*

$$d(S_t^{(\alpha)}(\mathbf{w}), \sigma_\alpha) \leq C' e^{-\nu t}, \qquad \forall t \geq 0 \tag{5.6.10}$$

(iii) "tangency" and "flatness": $\forall\, j = 1, \ldots, r$,

$$|\varphi^{(j)}(x_1, \ldots, x_r)| \leq C\,(x_1^2 + \ldots + x_r^2)^{\frac{3}{4}}. \tag{5.6.11}$$

Observations.
(1) The reader may be surprised by the fact that, for the first time in this book, an important property is appearing and being considered in class $C^{(k)}$ rather than in class C^∞: the reason is due to the fact that in Proposition 11 one cannot choose $k = +\infty$. In fact using the methods of Problems 3 and 4, p.428, the reader will check that in the equation $\dot{x} = \alpha x - x^3$, $\dot{z} = -z + x^2$ the surface σ_α cannot be of class $C^{(k)}$ for $k > \frac{1}{2\alpha}$.
(2) The above proposition is an important part of the "Hadamard and Perron theorem". It is sometimes called the "invariant" or "attractive manifold" theorem and it has importance in the development of the qualitative theory of differential equations. It has been intensely studied, undergoing many extensions and generalizations, often trivial but sometimes significant, [22].
(3) The family of surfaces σ_α is generally far from being uniquely determined by Eq. (5.6.1) (see the exercises for §5.6).
(4) The length of the proof and its formulae look quite discouraging. Actually the proof that follows is quite diluted and detailed (to conform to the spirit of this book). The subsections 5.6.A-5.6.D below have only a notational and definitorial character. The first technical step is in subsection 5.6.E with an application of the implicit function theorem with the purpose of stressing some properties of the surfaces $\sigma(\pi_t)$ approximating, as $t \to +\infty$, the surfaces that we are looking for. Subsection 5.6.F collects all the preceding inequalities to obtain further properties of the approximating surfaces $\sigma(\pi_t)$ for "very small" t. Furthermore, it contains the two basic ideas of the proof: (i) the estimates for very short times are possible because the quantity $\nu_0 = \min_{i=r+1,\ldots,d} -\mathcal{R}\mathrm{e}\,\lambda_i > 0$, measuring the attractivity of the stable directions, is much larger than all the other relevant quantities (i.e., for short times, the "strong attractivity of the stable directions prevails over the weak repulsivity of the unstable ones"); and (ii) the long-time estimates, as $t \to +\infty$, can be obtained from the ones for short times taking advantage of the autonomy of the equation.
These two themes occur again in a more or less repetitive way in subsections 5.6.G-5.6.N, all very similar to each other and which have been included here only for completeness.
The formulae are quite long and they could certainly be simplified and written more compactly. However, they are obtained by applying the procedures suggested in the text and they are left in the form in which they are constructed: in this way, the reader may easily recognize their various parts and their origin and this, perhaps, makes the proof more clear.
The vague attractivity assumption is used at the beginning of the proof to reduce it to an equivalent problem.
The proof says much more than what is stated in Proposition 11 and some of its corollaries are described in the problems at the end of the section.
The proof is adapted from that of Lanford.[12]

[12] See [29]

Proof. We shall discuss the proof of this proposition in the apparently particular case when $d = 2, r = 1$, and the equation is

$$\dot{x} = \alpha x + P(x, z), \quad \dot{z} = -\nu_0 z + Q(x, z), \tag{5.6.12}$$

where $\nu_0 > 0$ and P, Q are two $C^\infty(\mathcal{R}^2)$ functions with a second-order zero at the origin:

$$\begin{aligned} P(x, z) =& x^2\overline{P}_1(x, z) + z^2\overline{P}_2(x, z) + xz\overline{P}_3(x, z) \\ Q(x, z) =& x^2\overline{Q}_1(x, z) + z^2\overline{Q}_2(x, z) + xz\overline{Q}_3(x, z) \end{aligned} \tag{5.6.13}$$

and $\overline{P}_i, \overline{Q}_i$, $i = 1, 2, 3$, are in $C^\infty(\mathcal{R}^2)$, see Appendix B.

The stability matrix of $\mathbf{0} \in \mathcal{R}^2$ is

$$L_\alpha = \begin{pmatrix} \alpha & 0 \\ 0 & -\nu_0 \end{pmatrix}, \tag{5.6.14}$$

and we suppose that $\mathbf{x}_0 = \mathbf{0}$ is vaguely attractive near $\alpha_c = 0$.

This case looks quite special; however, its theory forces us to deal with all the difficulties of the general problem whose analysis is a repetition of that relative to Eq. (5.6.12). In the following formulae, it will essentially suffice to think that x and z are vectors with r and $d - r$ components and that α, ν_0 are matrices $r \times r$ or $(d - r) \times (d - r)$, respectively, possibly functions of the parameter α. The first will be a matrix with eigenvalues all having real part not less than $-\nu_0' > -\nu_0$, $\forall \in I = (\alpha_c - a, \alpha_c + a)$ and vanishing for $\alpha = \alpha_c$ and the second with all the eigenvalues with real part not exceeding $-\nu_0 < 0$, $\forall \alpha \in I$. Furthermore, P, Q also will have to be thought of as depending (smoothly) on α.

Hence, consideration of Eq. (5.6.12) does not diminish the real difficulties of the problem and treating it avoids puzzling with fictitious (mainly notational) difficulties the reader in his first approach to a proof which is complex, although quite natural in its development.

The interested reader will not have difficulties, on a second reading, in interpreting (*mutatis mutandis*) the proofs as relative to the general case (see exercises and problems at the end of this section for hints and suggestions).

To make the analysis of the proof easier, it will be divided it in various basic steps distinguished by alphabetic characters.

5.6.1 A: Preliminary Considerations and an Equivalent Problem.

Consider Eq. (5.6.12) and let U be the neighborhood introduced in Eq. (5.5.7), whose existence is guaranteed by the vague-attractivity assumption.

Let $\Gamma(\varrho) = \{$square in $\mathcal{R}^2$ with side size 2ϱ, centered at the origin$\} = \{\mathbf{w} \,|\, \mathbf{w} \in \mathcal{R}^2,\ |w_i| < \varrho, i = 1, 2\}$.

Choose $k = 0$, first. The case $k > 0$ will be discussed later. Fix $C \in (0, 1)$.

Let $\delta, a_+, t_0 \in (0,1)$ be small enough so that $a_+ < \frac{1}{2}a$ and the inequalities in Eqs. (5.6.41), (5.6.42), (5.6.43), footnote 13 in p.418, (5.6.53), (5.6.61), (5.6.76), (5.6.83), and (5.6.84) that will be met in the following discussion are satisfied. It is not worth listing them explicitly a priori here. The only fact that we shall really need is that they can all be simultaneously satisfied by choosing δ, a_+, t_0 small enough, once $C < 1$ is given.

Without loss of generality, we also suppose (see Definition 5, p.390, §5.5) that for all $\alpha \in I = (\alpha_c - a, \alpha_c + a)$,

$$\begin{aligned} &\Gamma(\delta) \subset U, \quad S_t^{(\alpha)} U \subset \Gamma(\frac{\delta}{2}), \quad \forall\, t \geq t_\delta, \\ &S_t^{(\alpha)} \Gamma(\delta_0) \subset \Gamma(\frac{\delta}{2}), \qquad\qquad \forall\, t \geq 0 \end{aligned} \tag{5.6.15}$$

for a suitable choice of $t_\delta > 0$ and of $\delta_0 < \delta$.

Let χ_δ be a $C^\infty(\mathcal{R}^2)$ function which takes values between 0 and 1, and has value 1 on $\Gamma(\frac{1}{2}\delta)$ and value 0 outside $\Gamma(\frac{2}{3}\delta)$. Let χ_δ have the form

$$\chi_\delta(x,z) = \chi(\frac{x}{\delta}, \frac{z}{\delta}), \tag{5.6.16}$$

where $\chi \in C^\infty(\mathcal{R}^2)$ is 1 on $\Gamma(\frac{1}{2})$ and 0 outside $\Gamma(\frac{2}{3})$.

So every motion beginning in U enters $\Gamma(\frac{1}{2}\delta)$, for good, in a finite time t_δ, and every motion beginning in $\Gamma(\delta_0$ never leaves $\Gamma(\frac{1}{2}\delta)$. This is a consequence of the vague-attractivity assumption.

It will then suffice to prove Proposition 11 for the equations

$$\dot{x} = \chi_\delta(x,z)\,(\alpha x + P(x,z)) \stackrel{def}{=} X_\delta(x,z,\alpha), \tag{5.6.17}$$

$$\dot{z} = -\nu_0 z + \chi_\delta(x,z)\,Q(x,z) = -\nu_0 + Z_\delta(x,z,\alpha), \tag{5.6.18}$$

It is useful to remark, for later use, that for the given values of $a_+, \delta, \delta_0, C, k$, and since χ_δ vanishes outside $\Gamma(\frac{2}{3}\delta)$, solutions $t \to S_t^{(\alpha,\delta)}(\mathbf{w})$ of Eqs. (5.6.17), (5.6.18) with initial datum $\mathbf{w} \in \Gamma(\delta)$ remain in $\Gamma(\delta)$: just note that

$$S_t^{(\alpha,\delta)}(x,z) = (x, z e^{-\nu_0 t}) \tag{5.6.19}$$

as long as $\chi_\delta(x, z e^{-\nu_0 t}) = 0$.

5.6.2 B: Some Useful Estimates of Derivatives.

Certain properties of solutions of Eqs. (5.6.17) and (5.6.18), thought of as an equation depending on the parameters α and δ and with datum $\mathbf{w} \in \Gamma(\delta)$ will be needed. The properties are summarized as follows. There exists a constant $M > 1$ and $t_0 \in (0,1)$ such that, $\forall\, t \in [-t_0, t_0]$, $\forall\, \alpha \in (-a, a), \forall\, \delta \in (0,1)$, $\forall\, j = 1,2$,

$$\Big|\frac{\partial S_t^{(\alpha,\delta)}(w_1,w_2)_i}{\partial w_j} - e^{-\mu_i t}\delta_{ij}\Big| \leq M|t|(|\alpha| + \delta), \tag{5.6.20}$$

$$\Big|\frac{\partial S_t^{(\alpha,\delta)}(w_1,w_2)_i}{\partial\alpha}\Big| \le M(\delta|t|\delta_{i1}+\delta^2|t|^2\delta_{i2}), \tag{5.6.21}$$

$$\Big|\frac{\partial S_t^{(\alpha,\delta)}(w_1,w_2)_i}{\partial t}+\mu_i w_i\Big| \le M|t|\big((|\alpha|+\delta)\delta_{i1}+\delta\,\delta_{i2}\big), \tag{5.6.22}$$

where $\mu_1 = 0, \mu_2 = \nu_0$, and where we have set $(x,z) = \mathbf{w}$, $\mathbf{w} = (w_1,w_2)$, and have denoted the components of $S_t^{(\alpha,\delta)}(\mathbf{w})$ as $S_t^{(\alpha,\delta)}(\mathbf{w})_i, i = 1,2$. Such notations will be often used in the following.

The above inequalities follow from an analysis of the regularity theorem for differential equations, §2.4, and they will be left to the reader, except Eq. (5.6.20) which is proved, as an example, in Appendix L.

We shall also need the following estimates, consequences of the definitions in Eqs. (5.6.17),(5.6.18). Let $\mathbf{w}\in\Gamma(\delta)$, $|\alpha|<a_+,\delta<1$, $i=1,2$; then

$$\Big|\frac{\partial X_\delta(\mathbf{w},\alpha)}{\partial w_i}\Big| \le M\,(|\alpha|+\delta), \qquad \Big|\frac{\partial X_\delta(\mathbf{w},\alpha)}{\partial\alpha}\Big| \le M\,\delta, \tag{5.6.23}$$

$$\Big|\frac{\partial Z_\delta(\mathbf{w},\alpha)}{\partial w_i}\Big| \le M\,\delta, \qquad \Big|\frac{\partial Z_\delta(\mathbf{w},\alpha)}{\partial\alpha}\Big| = 0, \qquad \text{and} \tag{5.6.24}$$

$$|Z_\delta(\mathbf{w},\alpha) \le M\,|\mathbf{w}|^2, \qquad |X_\delta(\mathbf{w},\alpha)| \le M\,(|\alpha||w_1|+|\mathbf{w}|^2), \tag{5.6.25}$$

where M can and will be chosen the same as before, possibly increasing the latter.

5.6.3 C: Definition of the Approximate Surfaces.

Let $\pi\in C^\infty(\mathcal{R})$ be such that, $\forall\, x\in[-\delta,\delta]$, it is $|\pi(x)|\le\delta$. Interpret it as defining a surface (a curve in this case, actually) $\sigma(\pi)\subset\Gamma(\delta)$ of parametric equations

$$z=\pi(x), \qquad x\in[-\delta,\delta]. \qquad \text{Also suppose that} \tag{5.6.26}$$

$$\Big|\frac{\partial\pi}{\partial x}\Big| \le C\sqrt{\delta}, \qquad x\in[-\delta,\delta] \tag{5.6.27}$$

(this choice of a bound on $\frac{\partial\pi(x)}{\partial x}$ is quite arbitrary: $C\sqrt{\delta}$ could equally well be replaced by $C\delta^\beta, 0<\beta<1$).

Then by the invariance of $\Gamma(\delta)$, the set $S_t^{(\alpha,\delta)}(\sigma(\pi))$, $t\ge 0$, is contained in $\Gamma(\delta)$ and, as will be seen shortly, it is a surface of the form $\sigma(\pi_t)$, where π_t is a new function verifying Eq. (5.6.27) and $|\pi_t|<\delta$.

It is then natural to try to define the surface that we are looking for as the surface $\sigma(\pi_\infty)$, where

$$\pi_\infty = \lim_{t\to+\infty}\pi_t \tag{5.6.28}$$

if this limit exists. In this case, in fact, the relation $S_t^{(\alpha,\delta)}(\sigma(\pi_\infty)) = \sigma(\pi_\infty)$ will be formally true.

5.6.4 D: Proof that the Approximate Surfaces are Well Defined.

First we look for an expression for π_t. This function should be defined by

$$(x, \pi_t) = S_t^{(\alpha,\delta)}(x_0, \pi(x_0)), \tag{5.6.29}$$

where x_0 is a suitable point in $[-\delta, \delta]$ defined, naturally, by Eq. (5.6.29) which should be thought of as an equation defining π_t and x_0 in terms of x and π. Such an equation certainly has a solution since

$$S_t^{(\alpha,\delta)}(\pm\delta, \pi(\pm\delta)) = (\pm\delta, \pi(\pm\delta)e^{-\nu_0 t}) \tag{5.6.30}$$

and, therefore, by continuity, there exists a "function" $A(x, t, \alpha, \pi)$ such that the abscissa of $S_t^{(\alpha,\delta)}(A(x, t, \alpha, \pi), \pi(A(x, t, \alpha, \pi)))$ is just x, i.e.,

$$x_0 = A(x, t, \alpha, \pi) \tag{5.6.31}$$

is the solution of the first equation obtained by equations the first component of Eq. (5.6.29). Then $\pi_t(x)$ can be defined as the second coordinate of $S_t^{(\alpha,\delta)}(x_0, \pi(x_0)$ with x_0 given by Eq. (5.6.31).

By Eq. (5.6.19), one naturally sets $A(x, t, \alpha, \pi) \equiv x$ for $|x| \geq \delta$.

It is not immediately clear from the above argument that the functions A and π are uniquely defined. To this question we devote the next step.

5.6.5 E: Alternative Proof of the Existence of π_t: Its Uniqueness for t Small and Estimates of Its Derivatives for t Small.

As already noted, the argument in subsection 5.6.D does not prove uniqueness of π_t nor does it allow to estimate its x derivative when one tries to check if it still verifies an inequality of the type of Eq. (5.6.27). In fact, it is a superfluous argument introduced just to help the reader to visualize what is done below.

It is possible to prove constructively the existence and uniqueness of the function A and, at the same time, to obtain an estimate of the derivatives of A with respect to x, t, α by using the implicit function theorem. To study the function A in this way, write Eq. (5.6.29) as

$$x = x_0 + \int_0^t X_\delta(S_t^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha)d\tau, \tag{5.6.32}$$

$$\pi_t(x) = \pi(x_0) + \int_0^t e^{-\nu_0(t-\tau)} Z_\delta(S_t^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha)d\tau, \tag{5.6.33}$$

obtained from Eqs. (5.6.17) and (5.6.18), pretending that X_δ and Z_δ are "known functions" of t and thinking of them as linear equations. We write Eq. (5.6.32) in the form $G_\pi(x, x_0, \alpha, t) = 0$, where

$$G_\pi(x, x_0, \alpha, t) = x_0 + \int_0^t X_\delta(S_t^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha)d\tau, \tag{5.6.34}$$

is a function in $C^\infty(\mathcal{R}^4)$ which will be mainly considered for $|x| \le \delta, |\alpha| \le 2a_+, |x_0| \le \delta, |t| \le t_0$.

We regard $G_\pi(x, x_0, \alpha, t) = 0$ as an equation for x_0 parameterized by x, α, t at fixed π.

Since the point $(\overline{x}, \overline{x}, \overline{\alpha}, 0)$ is a solution point of our equation, $\forall \overline{x} \in [-\delta, \delta]$, $\forall \overline{\alpha}, |\overline{\alpha}| < 2a_+$, we apply the implicit function theorem, see Appendix G, Eq. (G10), to find a square neighborhood with side $\varrho(\overline{x}, \overline{\alpha})$ of $(\overline{x}, \overline{\alpha}, 0)$ in $\mathcal{R}^3$ such that if

$$|x - \overline{x}|, |\alpha - \overline{\alpha}|, |t| \; < \varrho(\overline{x}, \overline{\alpha}) \tag{5.6.35}$$

then $G_\pi(x, x_0, \alpha, t) = 0$ has a solution $x_0 \in [-\delta, \delta]$.

To prove the existence of $\varrho(\overline{x}, \overline{\alpha})$, we must study the derivative

$$\frac{\partial G_\pi}{\partial x_0}(x, x_0, \alpha, t). \tag{5.6.36}$$

From Eq. (5.6.34), using Eqs. (5.6.20), (5.6.23), (5.6.26), (5.6.27), and also recalling that $C < 1$, $\delta < 1$ (so that $C\sqrt{\delta} < 1$), one finds

$$\begin{aligned}&\Big|\frac{\partial G_\pi}{\partial x_0}(x, x_0, \alpha, t) - 1\Big| = \Big|\int_0^t \frac{\partial X_\delta(S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha)}{\partial x_0}\, d\tau \\ &\le 8M\,(|\alpha| + \delta)(1 + M|t|(|\alpha| + \delta))\,|t|. \qquad \text{Furthermore}\end{aligned} \tag{5.6.37}$$

$$\frac{\partial G_\pi}{\partial x}(x, x_0, \alpha, t) \equiv 1, \tag{5.6.38}$$

and, setting $\boldsymbol{\xi}(t) \equiv \boldsymbol{\xi}(\alpha, \delta, x_0, \alpha, t) \stackrel{def}{=} S_t^{(\alpha,\delta)}(x_0, \pi(x_0))$, to simplify notations,

$$\begin{aligned}&\Big|\frac{\partial G_\pi}{\partial \alpha}(x, x_0, \alpha, t)\Big| = \Big|\int_0^t \Big(\big(\sum_{i=1}^2 \frac{\partial X_\delta(\boldsymbol{\xi}(\tau), \alpha)}{\partial w_i}\frac{\partial \xi(\tau)_i}{\partial \alpha}\big) + \frac{\partial X_\delta(\boldsymbol{\xi}(\tau), \alpha)}{\partial \alpha}\Big) d\tau \\ &\le |t|(|\alpha| + \delta)\delta|t|M^2 + \delta^2|t|^2M^2(|\alpha| + \delta + M\delta), \qquad \text{and, finally,}\end{aligned} \tag{5.6.39}$$

$$\Big|\frac{\partial G_\pi}{\partial t}(x, x_0, \alpha, t)\Big| = |X_\delta(\boldsymbol{\xi}(\alpha, \delta, x_0, \alpha, \tau))| \le 2M\delta(\delta + |\alpha|). \tag{5.6.40}$$

The above inequalities for the derivatives are valid for all $|t| \le 1, |x| \le d, |x_0| \le \delta$, $\delta \le 1$. Assume now a_+, δ, t_0 so small that $\forall\, |\alpha| < 2a_+$, $|t| \le t_0$:

$$\Big|\frac{\partial G_\pi}{\partial x_0}(x, x_0, \alpha, t) - 1\Big| < 10\, M\,(2a_+ + \delta)|t| < \frac{1}{2}, \tag{5.6.41}$$

$$\Big|\frac{\partial G_\pi}{\partial \alpha}(x, x_0, \alpha, t)\Big| \le 2\, M\, \delta\, |t| < \frac{1}{10}, \tag{5.6.42}$$

$$\Big|\frac{\partial G_\pi}{\partial t}(x, x_0, \alpha, t)\Big| \le M\, \delta\, (\delta + 2a_+) < \frac{1}{10}. \tag{5.6.43}$$

Here $\frac{1}{2}$ and $\frac{1}{10}$ are arbitrary small numbers, convenient for the upcoming estimates. Then if $|\overline{\alpha}| < a_+, |x-\overline{x}| < \frac{3}{4}\delta$, and if $\zeta = \min(a_+, t_0\frac{1}{4}\delta)$ the $\varrho(\overline{x},\overline{\alpha})$ just considered can be taken (see Appendix G, Proposition 1) as

$$\varrho(\overline{x},\overline{\alpha}) = \frac{\sigma}{2}\frac{\min\left|\frac{\partial G_\pi}{\partial x_0}\right|}{\max\left(\left|\frac{\partial G_\pi}{\partial x_0}\right| + \left|\frac{\partial G_\pi}{\partial x}\right| + \left|\frac{\partial G_\pi}{\partial \alpha}\right| + \left|\frac{\partial G_\pi}{\partial t}\right|\right)} \geq \frac{\zeta}{10} \tag{5.6.44}$$

having used Eqs. (5.6.41)-(5.6.43) to get the right-hand side inequality and having considered the maxima and the minima with respect to the parameters t, α, x, x_0 as they vary in $[-t_0, t_0], [-2a_+, 2a_+], [-\delta, \delta], [-\delta, \delta]$.

This shows the existence of A as a function of x, α, t as they vary in[13]

$$|x| \leq \delta, \quad |\alpha| \leq a_+, \quad |t| \leq t_+ \stackrel{def}{=} \frac{\zeta}{11} \tag{5.6.45}$$

and shows, as well, the possibility of estimating the derivatives of A as follows (see Eqs. (5.6.41)-(5.6.43) right-hand sides and Appendix G, Proposition 1):

$$\left|\frac{\partial A(x,t,\alpha,\pi)}{\partial x} - 1\right| \equiv \left| - \frac{\frac{\partial G_\pi(x,x_0,\alpha,t)}{\partial x}}{\frac{\partial G_\pi(x,x_0,\alpha,t)}{\partial x_0}} - 1\right| \leq 20\, M\,|t|(2a_+ + \delta), \tag{5.6.46}$$

$$\left|\frac{\partial A(x,t,\alpha,\pi)}{\partial \alpha}\right| \equiv \left| - \frac{\frac{\partial G_\pi(x,x_0,\alpha,t)}{\partial \alpha}}{\frac{\partial G_\pi(x,x_0,\alpha,t)}{\partial x_0}}\right| \leq 4\, M\,|t|\,\delta, \tag{5.6.47}$$

$$\left|\frac{\partial A(x,t,\alpha,\pi)}{\partial t}\right| \equiv \left| - \frac{\frac{\partial G_\pi(x,x_0,\alpha,t)}{\partial t}}{\frac{\partial G_\pi(x,x_0,\alpha,t)}{\partial x_0}}\right| \leq 4\, M\,\delta\,(2a_+ + \delta), \tag{5.6.48}$$

valid for x, α, t in the region of Eq. (5.6.45). It is important to stress that Eqs. (5.6.46)-(5.6.48) have been obtained independently of the choice of π provided

---

[13] Note that, if $2t_0M\,(a_+ + \delta)\delta < \min(a_+, t_0, \frac{\delta}{4}) = \zeta$, for $|x| \geq \frac{2}{3}\delta$ the determination of A is trivial and $A(x,t,\alpha,\pi) \equiv x$. Then let $\overline{x} \in [-\frac{3}{4}\delta, \frac{3}{4}\delta]$, $\overline{\alpha} \in [-a_+, a_+]$ and remark that by Eqs. (5.6.35) and (5.6.44) it is possible to solve uniquely the equation for $A \in [-\zeta, \zeta]$ in the region $|x-\overline{x}| < \frac{\zeta}{10}$, $|\alpha - \overline{\alpha}| < \frac{\zeta}{10}$, $|t| < \frac{\zeta}{10}$. As $\overline{x}, \overline{\alpha}$ vary in $[-\frac{3}{4}\delta, \frac{3}{4}\delta] \times [-a_+, a_+]$, this parallelepipedal region covers at least a neighborhood V of $[-\frac{2}{3}\delta, \frac{2}{3}\delta] \times [-a_+, a_+] \times [-\frac{\zeta}{11}, \frac{\zeta}{11}]$.

By the uniqueness of A in each parallelepiped, the functions A thus defined coincide at the points which are common to several parallelepipeds. Furthermore, the functions A have a value equal to x for $|x| > \frac{2}{3}\delta$.

Hence, we have built a continuous piecewise-differentiable solution A of $G_\pi(x, A, \alpha, t) = 0$ in the region of Eq. (5.6.45); and by construction, A is the unique solution, with the property $|A - x| < \zeta$, in this region.

Actually, A must be C^∞ in the region of Eq. (5.6.45), since in each of the parallelepipeds where A has been constructed, A has this property and we have uniqueness.

Finally A is the only solution with $|A| < \delta$ because, as noted above, any such solution must verify $|A - x| < \zeta < \frac{1}{4}\delta$ and for $|x| \geq \frac{2}{3}\delta$ it is $A \equiv x$.

$$|\pi(x)| \le \delta, \qquad \text{and} \qquad \Big|\frac{\partial \pi(x)}{\partial x} \le C\sqrt{\delta}, \qquad \forall\, x \in [-\delta, \delta]. \tag{5.6.49}$$

The above considerations show that the function π_t is well defined at least for $|\alpha| < a_+, |t| < t_+$, via Eqs. (5.6.29) and (5.6.31).

The uniqueness of the A function, coming from its construction (see footnote 13, p. 417) allows us to conclude that π_t is the $S_t^{(\alpha,\delta)}$ image of $\sigma(\pi)$: $S_t^{(\alpha,\delta)}\sigma(\pi) = \sigma(\pi_t)$. Also note that, by the invariance of $\Gamma(\delta)$, one has $S_t^{(\alpha,\delta)}\sigma(\pi) \subset \Gamma(\delta)$.

The invariance of $\Gamma(\delta)$ for the motions generated by Eqs. (5.6.17) and (5.6.18) also implies that π_t, verifies the first of Eqs. (5.6.49) (a property already encountered during the construction of A).

5.6.6 F: Check of the Validity of Eq. (5.6.49) for π_t, $0 \le t \le t_+$

This check is of fundamental importance since it will allow us to define π_t for all $t \ge 0$.

The relation $S_t^{(\alpha,\delta)}\sigma(\pi) = \sigma(\pi_t)$, $t \in [0, t_+]$ will guarantee, taking also into account the group property $S_t^{(\alpha,\delta)}S_{t'}^{(\alpha,\delta)} = S_{t+t'}^{(\alpha,\delta)}$, that if $t \in [0, t_+)$, $t' \in [0, t_+), t + t' \in [0, t_+]$ and if $\pi_t, \pi_{t'}, \pi_{t+t'}$ verify Eq. (5.6.49), then

$$(\pi_t)_{t'} = \pi_{t+t'}. \tag{5.6.50}$$

This relation will allow us to define uniquely π_t, $\forall\, t \ge 0$, by dividing the interval $[0, t]$ into intervals with amplitude $\tau < t_+$ and, then, recursively setting

$$\pi_t = (\pi_{t-\tau})_\tau = ((\pi_{t-2\tau})_\tau)_\tau. \tag{5.6.51}$$

The definition will necessarily coincide with the one that could be given by setting $S_t^{(\alpha,\delta)}\sigma(\pi) = \sigma(\pi_t)$, $t \ge 0$.

Therefore let us verify that, if $0 \le t \le t_+$, π_t fulfills the second of Eqs. (5.6.49) (as noted above, the first has already been checked).

For this purpose, we use Eq. (5.6.33), where instead of x_0, one should imagine $A(x, t, \alpha, \pi)$. Differentiating both sides, one finds

$$\begin{aligned}\frac{\partial \pi_t}{\partial x} &= \frac{\partial \pi_t(x_0)}{\partial x_0}\frac{\partial A}{\partial x}e^{-\nu_0 t} + \sum_{i=1}^{2}\int_0^t \Big[e^{-\nu_0(t-\tau)}\frac{\partial Z_\delta}{\partial w_i}(S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha)\\ &\times \Big\{\frac{\partial S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0))_i}{\partial x_0} + \frac{\partial S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0))_i}{\partial \pi(x_0)}\frac{\partial \pi(x_0)}{\partial x_0}\Big\}\frac{\partial A(x, \tau, \alpha, \pi)}{\partial x}\Big]\, d\tau\end{aligned} \tag{5.6.52}$$

with slightly symbolic differentiation notations (hopefully self-explanatory). By Eqs. (5.6.41), (5.6.46), (5.6.49), and (5.6.20), (5.6.24), Eq. (5.6.52) implies, with some labor, that $\forall\, t \in [0, t_+]$, $\forall\, \alpha \in [-a_+, a_+]$, $\forall\, x \in [-\delta, \delta]$,

$$
\begin{aligned}
&\Big|\frac{\partial \pi_t(x)}{\partial x}\Big| \le C\sqrt{\delta}\, e^{-\nu_0 t}(1 + 20\, M\,(2a_+ + \delta)t)\\
&+ tM\delta\Big\{(1 + Mt(a_+ + \delta)) + M\,t\,(a_+ + \delta)C\sqrt{\delta} + M\,t\,(a_+ + \delta)\Big\}\\
&= C\sqrt{\delta}\,(1 + Mt\,(a_+ + \delta))\cdot(1 + 2 -\ M\,(2a_+ + \delta)t)\\
&+ C\sqrt{\delta}\Big\{e^{-\nu_0 t}(1 + 20\, M\, t(2a_+ + \delta)) + t\, M\, C^{-1}\sqrt{\delta}\\
&\times\Big[(1 + M\,t\,(a_+ + \delta)) + M\,t\,(a_+ + \delta)C\sqrt{\delta} + M\,t\,(a_+ + \delta)\\
&+ C\sqrt{\delta}(1 + M\,t\,(a_+ + \delta))\Big](1 + 20M\,(2a_+ + \delta))\\
&\times(1 + 20M\,(2a_+ + \delta))\Big\} \le C\sqrt{\delta}\,(1 - \frac{\nu_0 t}{2})
\end{aligned}
\tag{5.6.53}
$$

if δ, a_+, t_0 (recall that $t_+ \le t_0$) are supposed to have been so chosen that the last inequality in Eq. (5.6.53) holds, $\forall\ t \in [0, t_+]$.[14]

The above arguments prove that π_t can be defined by $S_t^{(\alpha,\delta)}\sigma(\pi) = \sigma(\pi_t)$ or, equivalently, by Eq. (5.6.51), for $t \ge 0$ and show that π_t verifies Eq. (5.6.49) for all $t \ge 0$.

5.6.7 G: Proof of the Existence of the Limit as $t \to +\infty$ of π_{nt} for $t \in [0, t_+]$.

We shall proceed by recursively evaluating

$$||\pi_{nt} - \pi_{(n-1)t}|| = \max_{|x|\le\delta} |\pi_{nt} - \pi_{(n-1)t}(x)| \tag{5.6.54}$$

and show that the series

$$\sum_{n=0}^{\infty} ||\pi_{nt} - \pi_{(n-1)t}|| < +\infty \tag{5.6.55}$$

converges. This implies that π_{nt} converges uniformly as $n \to +\infty$ to a limit.

To study the series of Eq. (5.6.55), consider two functions π, π' verifying Eq. (5.6.49) and, through them, construct the functions $A(x, t, \alpha, \pi)$ and $A(x, t, \alpha, \pi')$ defined on the set given by Eq. (5.6.45), solving the equations for x_0: $G_\pi(x, x_0, t, a) = 0$ and $G_{\pi'}(x, x_0, t, \alpha) = 0$ as indicated in subsection 5.6.E.

Shortening $A(x, t, \alpha, \pi)$ and $A(x, t, \alpha, \pi')$ in x_0, x_0', respectively, and using Eq. (5.6.33), one then has $\forall\, t \in [0, t_+]$,

[14] One sees that $C\sqrt{\delta}$ could be replaced by $C\delta^\gamma$, $\gamma < 1$. The choice $\gamma = 1$ could only be made if ν_0 is large enough (or if we decided to allow $C > 1$ and C to be large enough.)

$$|\pi_t(x) - \pi'_t(x)| \le e^{-\nu_0 t}|\pi(x_0) - \pi'(x'_0)| + \int_0^t d\tau e^{-\nu_0(t-\tau)} \cdot |Z_\delta(S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha) - Z_\delta(S_\tau^{(\alpha,\delta)}(x'_0, \pi'(x'_0)), \alpha)| \tag{5.6.56}$$

which, by Eqs. (5.6.24), (5.6.49), and (5.6.20), implies

$$\begin{aligned}|\pi_t(x) - \pi'_t(x)| \le & e^{-\nu_0 t}(|\pi(x_0) - \pi'(x'_0)| + |\pi'(x_0) - \pi'(x'_0)|) \\ & + \int_0^t M\delta \sum_{i=1}^{2} |S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0))_i - S_\tau^{(\alpha,\delta)}(x'_0, \pi'(x'_0))_i|\, d\tau \\ \le & e^{-\nu_0 t}(||\pi - \pi'|| + C\sqrt{\delta}\, |x_0 - x'_0|) + \int_0^t d\tau \\ & 2M\delta(1 + M\delta(a_+ + \delta)\tau)(|x_0 - x'_0| + |\pi(x_0) - \pi'(x_0)|) \\ & \le ||\pi - \pi'||\, (e^{-\nu_0 t}| + 2M\delta t(1 + M\delta(a_+ + \delta)t)) \\ & + |x_0 - x'_0|(C\sqrt{\delta}e^{-\nu_0 t} + 2M\delta t(1 + M\delta(a_+ + \delta)t)).\end{aligned} \tag{5.6.57}$$

We must therefore estimate $|x_0 - x'_0|$. Remark that $(x_0, \pi(x_0))$ and $(x_0, \pi'(x_0))$ are the values of $S_\tau^{(\alpha,\delta)}(x, \pi_t(x))$ and $S_\tau^{(\alpha,\delta)}(x, \pi'_t(x))$ hence, as in Eq. (5.6.32),

$$\begin{aligned} x_0 =& x - \int_0^t d\tau\, X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi_t(x)), \alpha), \\ x'_0 =& x - \int_0^t d\tau\, X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi'_t(x)), \alpha),\end{aligned} \tag{5.6.58}$$

Then, by Eqs. (5.6.23) and (5.6.20),

$$\begin{aligned}|x_0 - x'_0| \le & \int_0^t d\tau \left| X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi_t(x)), \alpha) - X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi'_t(x)), \alpha)\right| \\ \le & \int_0^t d\tau\, M(a_+ + \delta)1(1 + M\tau(a_+ + \delta))|\pi_t(x) - \pi'_t(x)| \\ \le & t\, M(a_+ + \delta)1(1 + M\, t(a_+ + \delta))|\pi_t(x) - \pi'_t(x)|,\end{aligned} \tag{5.6.59}$$

Hence Eqs. (5.6.57) and (5.6.59) imply

$$\begin{aligned}|\pi_t(x) - \pi'_t(x)| \le & ||\pi - \pi'||(e^{-\nu_0 t} + 2M\delta t(1 + M\delta(a_+ + \delta)t)) \\ & + (C\sqrt{\delta}e^{-\nu_0 t} + 2M\delta t(1 + M\delta(a_+ + \delta)t)) \\ & \times (2M(a_+\delta)t(1 + Mt(a_+ + \delta)))|\pi_t(x) - \pi'_t(x)|.\end{aligned} \tag{5.6.60}$$

This formula implies a bound on $|\pi_t(x) - \pi'_t(x)|$ if a_+, δ, t_0 are so small that for all t, $0 \le t \le t_0$ holds the inequality

$$\begin{aligned}&\frac{e^{-\nu_0 t}+2M\delta t(1+M\delta(a_++\delta)t)}{1-2M(a_+\delta)t(1+Mt(a_++\delta))}\\&\times\Big(C\sqrt{\delta}e^{-\nu_0 t}+2M\delta t\,(1+M\delta(a_++\delta)t)\Big)\le(1-\frac{\nu_0 t}{2}).\end{aligned}\tag{5.6.61}$$

Equations (5.6.61) and (5.6.60) imply $|\pi_t(x)-\pi'_t(x)|\le(1-\frac{1}{2}\nu_0 t)||\pi-\pi'||$; hence, $\forall\,\alpha\in[-a_+,a_+]$, $\forall\,t\in[0,t_+]$,

$$||\pi_t-\pi'_t||\le(1-\frac{\nu_o t}{2})\,||\pi-\pi'||.\tag{5.6.62}$$

A similar calculation would allow us to show that if π verifies Eq. (5.6.49) and a_+,δ,t_0 are sufficiently small,

$$||\pi_t-\pi_{t'}||\le\gamma\,|t-t'|\tag{5.6.63}$$

for all $\alpha\in[-a_+,a_+]$, $\forall\,t,t'\in\mathcal{R}_+$, $|t-t'|<t_+$ provided γ is suitably chosen.

We shall use this inequality without proof here (see Appendix M where a proof is discussed and an explicit expression for γ is exhibited).

Equation (5.6.62) allows us to estimate recursively Eq. (5.6.54) since it holds under the sole assumption that π and π' verify Eq. (5.6.49) and $t\in[0,t_+]$, $\alpha\in[-a_+,a_+]$. By subsection 5.6.F, one finds

$$||\pi_{nt}-\pi_{(n-1)t}||\le\big(1-\frac{\nu_0 t}{2}\big)^{n-1}\,||\pi_t-\pi||\le2\delta\,\big(1-\frac{\nu_0 t}{2}\big)^{n-1}\tag{5.6.64}$$

valid for all π verifying Eq. (5.6.49), $\forall\,n$ integer and ≥1.

Hence, the series of Eq. (5.6.55) is uniformly convergent as π varies in the class of the functions verifying Eq. (5.6.49), $\forall\,t\in[0,t_+]$, $\forall\,\alpha\in[-a_+,a_+]$.

5.6.8 H: Independence of the Limit as $n\to+\infty$ of π_{nt} from π and $t\in[0,t_+]$

Denote π the continuous function defined on $[-\delta,\delta],\forall\,t\in[0,t_+]$, in terms of a π verifying Eq. (5.6.49), by

$$\lim_{n\to+\infty}\pi_{nt}=\pi_{\infty,t,\pi},\tag{5.6.65}$$

the continuity being insured by the uniformity of the limit of Eq. (5.6.65), see Eq. (5.6.55). The function $\pi_{\infty,t,\pi}$ is π independent. In fact, Eq. (5.6.62) recursively implies

$$||\pi_{nt}-\pi'_{nt}||\le(1-\frac{\nu_0 t}{2})^n||\pi-\pi'||\xrightarrow[n\to+\infty]{}0\tag{5.6.66}$$

if π,π' verify Eq. (5.6.49). Hence it will be simply denoted as π_t. Now let $t',t\in[0,t_+]$ and $t'/t=p/q=$ rational number, p,q integers, then

$$\pi_{ntp} = \pi_{nt'q}, \tag{5.6.67}$$

hence, in the limit $n \to +\infty$, Eq. (5.6.67) implies

$$\pi_{\infty,t} = \pi_{\infty,t'} \tag{5.6.68}$$

if $t/t' = \{\text{rational number}\}$. Therefore Eq. (5.6.63) implies that Eq. (5.6.68) holds for all $t, t' \in (0, t_+]$ and $\pi_{\infty,t}$ is t independent. Denoting π_∞ the function in Eq. (5.6.68), it is $(\pi_\infty)_t \equiv \pi_\infty$ and this proves the invariance of $\sigma(\pi_\infty)$; hence, Eq. (5.6.9).

5.6.9 I: Attractivity of $\sigma(\pi_\infty)$.

Given $(\overline{x}, \overline{z}) \in \Gamma(\delta)$, let $\overline{\pi}$ be a function verifying Eq. (5.6.49) and $\overline{\pi}(\overline{x}) = \overline{z}$, e.g., $\overline{\pi}(\overline{x}) = \overline{z}$, $x \in [-\delta\delta]$. Given $t > 2t_+$, let $\overline{t} \in (0, t_+)$ such that $\overline{t} > \frac{1}{2}t_+$ and, furthermore, $t/\overline{t} = N =$ integer. Then, by Eq. (5.6.66) or Eq. (5.6.62),

$$\begin{aligned} &||\overline{\pi}_t - \pi_\infty|| \equiv ||\overline{\pi}_t - (\pi_\infty)_t|| \\ &= ||(\overline{\pi})_{Nt} - (\pi_n)_{N\overline{t}}|| \le (1 - \frac{\nu_0 \overline{t}}{2})^N ||\pi - \pi_n|| \\ &\le (1 - \frac{\nu_0 \overline{t}}{2})^N ||\overline{\pi} - \pi_\infty|| \le 2\delta\, (1 - \frac{\nu_0 \overline{t}_+}{4})^{t/t_+} \end{aligned} \tag{5.6.69}$$

which proves that $\sigma(\overline{\pi}_t)$, hence $S_t^{(\alpha,\delta)}(\overline{x}, \overline{z})$ as well, approaches $\sigma(\pi_\infty)$ with exponential strength so that the attractivity of $\sigma(\pi_\infty)$ is proved in the case of the Eqs. (5.6.17) and (5.6.18) and this immediately leads to Eq. (5.6.10).

5.6.10 L: Order of Tangency.

Let us show that if π is chosen so that it verifies Eq. (5.6.49) as well as

$$|\pi(x)| \le C|x|^{\frac{3}{2}}, \qquad \forall\, x \in [-\delta, \delta], \tag{5.6.70}$$

then it is also true that

$$|\pi_t(x)| \le C|x|^{\frac{3}{2}}, \qquad \forall\, x \in [-\delta, \delta],\ \forall\, t \in \mathcal{R}_+. \tag{5.6.71}$$

Hence, for $x \in [-\delta, \delta]$, $|\pi_\infty(x)| \le C|x|^{\frac{3}{2}}$, implying Eq. (5.6.11) for $k = 0$.

Suppose that π verifies Eqs. (5.6.49) and (5.6.70), e.g., $\pi(x) \equiv \frac{2}{3}C|x|^{\frac{3}{2}}$. From Eqs. (5.6.33), (5.6.25),(5.6.20) and $S_t^{(\alpha,\delta)}(0,0) \equiv (0,0)$, it follows

$$\begin{aligned} |\pi_t(x)| &\le e^{-\nu_0 t} C|x_0|^{\frac{3}{2}} + \int_0^t M\, |S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0))|^2 d\tau \\ &\le e^{-\nu_0 t} C|x_0|^{\frac{3}{2}} + Mt(1 + 2Mt(a_+ + \delta))^2 (x_0^2 + \pi(x_0)^2) \\ &\le C|x_0|^{\frac{3}{2}} (e^{-\nu_0 t} C + Mt(1 + 2Mt(a_+ + \delta))^2 (C^{-1} \sqrt{|x_0|} + C|x_0|^{\frac{3}{2}})) \end{aligned} \tag{5.6.72}$$

for $t \in [0, t_+]$, $\alpha \in [-a_+, a_+]$.
From Eqs. (5.6.68),(5.6.25),(5.6.20) and $S_f^{(\alpha,\delta)}(0,0) \equiv (0,0)$

$$\begin{aligned}|x_0| \leq & |x| + M \int_0^t (a_+ |S_{-\tau}^{(\alpha,\delta)}(x, \pi_t(x))_1| + |S_{-\tau}^{(\alpha,\delta)}(x, \pi_t(x))|^2)\, d\tau \\ \leq & |x| + Mt\big\{a_+[(1 + Mt(a_+ + \delta))(|x| + |\pi_t(x))|] \\ & + [2(1 + 2Mt(a_+ + \delta))^2(|x|^2 + |\pi_t(x)|^2)]\big\} \\ \leq & |x|\Big(1 + Mt\big\{a_+[(1 + Mt(a_+ + \delta))] + [2(1 + 2Mt(a_+ + \delta))^2\delta]\big\}\Big) \\ & + |\pi_t(x)| Mt\Big\{a_+[(1 + Mt(a_+ + \delta))] + [2(1 + 2Mt(a_+ + \delta))^2\delta]\Big\}.\end{aligned} \tag{5.6.73}$$

To simplify the notations, rewrite Eqs. (5.6.72) and (5.6.73) by observing that if $a_+, \delta, t_+ < 1$ (as supposed since the beginning of the analysis), there exists $M' > 0$ such that

$$|\pi_t(x)| \leq C|x|^{\frac{3}{2}}\big(1 - \frac{\nu_0 t}{2} + M'\sqrt{\delta}\, t\big) \tag{5.6.72'}$$

$$|x_0| \leq |x|\,\big((1 + M'(a_+ + \delta)t\big) + |\pi_t(x)|\,\big((1 + M'(a_+ + \delta)t\big) \tag{5.6.73'}$$

Then, taking the $\frac{2}{3}$ power of Eq. (5.6.72') and using and (5.6.73')

$$|\pi_t(x)|^{\frac{3}{2}} \leq C^{\frac{2}{3}}\big(1 - \frac{\nu_0 t}{2} + M'\sqrt{\delta}\, t\big)\big(1 + M'(a_+ + \delta)t\big)(|x| + |\pi_t(x)|), \tag{5.6.74}$$

Since $\delta < 1$, $|\pi_t(x)| \leq \delta$, using $|\pi_t(x)| \leq |\pi_t(x)^{\frac{3}{2}}$ deduce from Eq. (5.6.74)

$$|\pi_t(x)|^{\frac{2}{3}} \leq C^{\frac{2}{3}}|x| \frac{(1 - \frac{\nu_0 t}{2} + M'\sqrt{\delta}\, t)^{\frac{2}{3}}(1 + M'(a_+ + \delta)t)}{1 - M'(a_+ + \delta)t(1 - \frac{\nu_0 t}{2} + M'\sqrt{\delta}\, t)^{\frac{2}{3}}} \tag{5.6.75}$$

Hence, let us choose φ, a_+, t_0 to be so small that, $\forall\, t \in [0, t_0]$, the ratio in Eq. (5.6.75) is bounded by

$$\text{ratio in Eq. (5.6.75)} \;\leq\; 1 - \frac{\nu_0 t}{4}, \tag{5.6.76}$$

we see that Eqs. (5.6.75) and (5.6.76) imply, $\forall\, x \in [-\delta, \delta]$, $\forall\, t \in [0, t_+]$,

$$|\pi_t(x)| \leq C|x|^{\frac{3}{2}}(1 - \frac{\nu_0 t}{4}) \leq C|x|^{\frac{3}{2}}, \tag{5.6.77}$$

hence, the inequality between the left-hand and the right-hand sides holds, $\forall\, t \geq 0$, and this implies Eq. (5.6.11) for $k = 0$.

5.6.11 M: Regularity in α.

This is the last property to check. One proceeds almost exactly in the same way as above. Details will be illustrated because in some sense there is here a technical idea, new with respect to the ones already met. Actually we shall prove that π is a Lipshitzian function of α and x for $\alpha \in [-a_+, a_+]$, $x \in [-\delta, \delta]$, i.e., a somewhat stronger result.

Consider a function $(x, \alpha) \to \pi(x, \alpha)$ defined for $x \in [-\delta, \delta]$, $\alpha \in [-a_+, a_+]$, of class $C^{(1)}$ and verifying Eq. (5.6.49) for each α. Define $\pi_t(x, \alpha)$ by thinking of π as a function of x for each α and proceeding as in subsection 5.6.E. From Eqs. (5.6.33), (5.6.24), (5.6.21), (5.6.20), (5.6.49) and employing the usual notations, one finds that

$$\pi_t(x,\alpha) = e^{-\nu_0 t}\pi(x_0,\alpha)\int_0^t e^{-\nu_0(t-\tau)} Z_\delta(S_\tau^{(\alpha,\delta)}(x_0,\pi(x_0,\alpha)),\alpha)\,d\tau, \quad (5.6.78)$$

hence, recalling that x_0 is also α dependent and denoting $\partial_\alpha \stackrel{def}{=} \frac{\partial}{\partial\alpha}$:

$$\begin{aligned}
&|\partial_\alpha \pi_t(x,\alpha)\Big| \le e^{-\nu_0 t}|\partial_\alpha\pi(x_0,\alpha) + \frac{\partial\pi(x_0,\alpha)}{\partial x}\frac{\partial x_0}{\partial\alpha}\Big| + \int_0^t d\tau\Big[\sum_{i=1}^2 \\
&\Big|\frac{\partial Z_\delta}{\partial w_i}(S_\tau^{(\alpha,\delta)}(x_0,\pi(x_0,\alpha)))\frac{d}{d\alpha}\{S_\tau^{(\alpha,\delta)}(x_0,\pi(x_0,\alpha))_i\}\Big|\Big]e^{-\nu_0(t-\tau)} \\
&\le e^{-\nu_0 t}|\partial_\alpha\pi(x_0,\alpha)| + e^{-\nu_0 t}C\sqrt{\delta}|\frac{\partial x_0}{\partial\alpha}| \\
&+ 2M\delta\int_0^t d\tau\Big\{M\tau(\delta+\delta^2 t) + (1+M\tau(a_+ +\delta)) \\
&\times\Big(|\frac{\partial x_0}{\partial\alpha}| + |\frac{\partial\pi(x_0,\alpha)}{\partial x_0}\frac{\partial x_0}{\partial\alpha}| + |\frac{\partial\pi(x_0,\alpha)}{\partial\alpha}|\Big)\Big\} \qquad (5.6.79)\\
&\le e^{-\nu_0 t}|\partial_\alpha\pi(x_0,\alpha)| + e^{-\nu_0 t}C\sqrt{\delta}|\frac{\partial x_0}{\partial\alpha}| + M^2\delta^2(1+\delta t)t^2 \\
&+ 2M\delta(1+M\tau(a_+ +\delta))t\big((1+C\sqrt{\delta})|\frac{\partial x_0}{\partial\alpha}| + |\frac{\partial\pi(x_0,\alpha)}{\partial\alpha}|\big) \\
&\le e^{-\nu_0 t}2M\delta(1+M\tau(a_+ +\delta))t|\frac{\partial\pi(x_0,\alpha)}{\partial\alpha}| + M\delta^2(1+\delta t)t^2 \\
&+ \big\{e^{-\nu_0 t}C\sqrt{\delta} + 2M\delta(1+Mt(a_+ +\delta)t(1+C\sqrt{\delta}))\big\}|\frac{\partial x_0}{\partial\alpha}|.
\end{aligned}$$

The $\partial_\alpha x_0$ is estimated as in subsection 5.6.G, by Eq. (5.6.58) rewritten as

$$x_0 = x - \int_0^t X_\delta(S_{-\tau}^{(\alpha,\delta)}(x,\pi_t(x,\alpha),\alpha)d\tau, \qquad (5.6.80)$$

hence, proceeding as in the derivation of Eq. (5.6.79) and using Eq. (5.6.23):

$$
\begin{aligned}
|\frac{\partial x_0}{\partial \alpha}| \leq & \int_0^t d\tau \Big\{ M\delta + 2M(a_+ + \delta) \\
& \cdot \Big[M\delta\tau(1+\delta\tau) + (1 + M\tau(a_+ + \delta) | \frac{\partial \pi_t(x,\alpha)}{\partial \alpha} |) \Big] \Big\} \\
\leq & M + 2M^2\delta \frac{t^2}{2}(1+)(a_+ + \delta) \\
& + t(1 + Mt(a_+ + \delta))2M(a_+ + \delta) | \frac{\partial \pi_t(x,\alpha)}{\partial \alpha} |.
\end{aligned}
\tag{5.6.81}
$$

Then Eqs. (5.6.79) and (5.6.81) imply

$$
|\partial_\alpha \pi_t(x,\alpha)| \leq \frac{A + B|\partial_\alpha \pi(x_0,\alpha)|}{1 - GS} \tag{5.6.82}
$$

with

$$
\begin{aligned}
G & \stackrel{def}{=} [e^{-\nu_0 t} C\sqrt{\delta} + 2M\delta(1 + Mt(a_+ + \delta))t(1 + C\sqrt{\delta})], \\
S & \stackrel{def}{=} t[1 + Mt(a_+ d)2M(a_+ + \delta)], \\
A & \stackrel{def}{=} M^2\delta^2(1+\delta t)t^2 + G[M + M^2\delta t^2(1+\delta t)(a_+ + \delta)], \\
B & \stackrel{def}{=} e^{-\nu_0 t} + 2M\delta(1 + Mt(a_+ + \delta))t
\end{aligned}
$$

and to understand the essential features of Eq. (5.6.82), we note that if δ, a_+, t_0 to are chosen so small that there is $\overline{M}$ such that the first term in Eq. (5.6.82) can be bounded by

$$
\frac{A}{1 - GS} \leq \overline{M}\delta\sqrt{\delta}\, t \tag{5.6.83}
$$

for all $t \in [0, t_0]$, $\forall\, \alpha \in [-a_+, a_+]$, and the coefficient of $|\partial_\alpha \pi(x_0,\alpha)|$ in Eq. (5.6.82) can be bounded as

$$
\frac{B}{1 - GS} \leq 1 - \frac{\nu_o t}{2}, \tag{5.6.84}
$$

then Eq. (5.6.82) can be simply rewritten, $\forall\, t \in [0, t_+]$, $\forall\, \alpha \in [-a_+, a_+]$,

$$
||\partial_\alpha \pi_t|| \leq \max_{\substack{|x|\leq\delta \\ |\alpha|\leq a_+}} |\partial_\alpha \pi_t(x,\alpha)| \leq \overline{M}\delta^{\frac{3}{2}} t + (1 - \frac{\nu_0 t}{2})||\partial_\alpha \pi|| \tag{5.6.85}
$$

Now fix π to be a function of the variable x only and verifying Eq. (5.6.49). Apply Eq. (5.6.85) to the functions $\pi_{nt}, \pi_{(n-1)t}, \ldots$ thought of as functions of x and α. If $t \in (0, t+]$, $n = 0, 1, 2, \ldots$,

$$
||\partial_\alpha \pi_{nt}|| \leq \overline{M}\delta^{\frac{3}{2}} t + (1 - \frac{\nu_0 t}{2})||\partial_\alpha \pi_{(n-1)t}||. \tag{5.6.86}
$$

Then, Eq. (5.6.86) implies, recursively,

$$||\partial_\alpha \pi_{nt}|| \leq \overline{M}\delta^{\frac{3}{2}} t(1 + (1 - \frac{\nu_0 t}{2}) + (1 - \frac{\nu_0 t}{2})^2 + \ldots) \tag{5.6.87}$$

because $\pi_0 = \pi$ is by hypothesis α independent, so that $\partial_\alpha \pi_0 \equiv 0$, i.e.,

$$||\partial_\alpha \pi_{nt}|| \leq \frac{\overline{M}\delta^{\frac{3}{2}} t}{\nu_0 t/2} \equiv \frac{2\overline{\delta}^{\frac{3}{2}}}{\nu_0}. \tag{5.6.88}$$

The regularity of π_∞ can now be checked:

$$\begin{aligned} &|\pi_\infty(x,\alpha) - \pi_\infty(x',\alpha')| = \lim_{n\to\infty} |\pi_{nt}(x,\alpha) - \pi_{nt}(x',\alpha')| \\ &\leq \lim_{n\to\infty} (|x - x'| + |\alpha - \alpha'|) \max \big(|\frac{\partial \pi_{nt}}{\partial x}| + |\frac{\partial \pi_{nt}}{\partial \alpha}|\big) \end{aligned} \tag{5.6.89}$$

where the maximum is taken on the set $[-\delta,\delta]\times[-a_+,a_+]$ and, by Eq. (5.6.49) (considered for π_{nt}) and Eq. (5.6.88), it can be estimated by $D = \sqrt{\delta}\,(1 + 2M\nu_0^{-1})$. Hence,

$$|\pi_\infty(x,\alpha) - \pi_\infty(x',\alpha')| \leq D(|x - x'| + |\alpha - \alpha'|), \tag{5.6.90}$$

showing that π_∞ is continuous in x and α (i.e., it is in class $C^{(0)}$) and, actually, that it is a Lipshitz function in x and α (with a Lipshitz constant D which can be taken as small as desired by taking δ small enough).

5.6.12 N: General Case.

To show that π_∞ is k-times differentiable with respect to x if a_+, δ are chosen sufficiently small, one proceeds to estimate $\frac{\partial^2 \pi_t}{\partial x^2}$ and, successively, $\frac{\partial^3 \pi_t}{\partial x^3}, \ldots, \frac{\partial^{k+1}\pi_t}{\partial x^{k+1}}$ in the same way as in the $k = 0$ case we studied π_t and $\frac{\partial \pi_t}{\partial x}$ to show that π_∞ was $C^{(0)}$, assuming now that π is in $C^{(k+1)}([-\delta,\delta])$ to start with.

Proceeding with the same technique as in subsections 5.6.F, 5.6.G, and 5.6.L, $(1 + k)\delta, (1 + k)a_+, t_0$ are chosen sufficiently small so that inequalities similar to Eqs. (5.6.41), (5.6.42), (5.6.43), and (5.6.53), etc. hold. One finds

$$||\frac{\partial^h \pi_t}{\partial x^h}|| \equiv \max_{|x|\leq\delta} |!\frac{\partial^h \pi_t(x)}{\partial x^h}|| \leq (1 - \frac{\nu_0 t}{2})||\frac{\partial^h \pi}{\partial x^h}|| + tR_{k,\delta}(\sum_{j=1}^{h-1} ||\frac{\partial^j \pi}{\partial x^j}||) \tag{5.6.91}$$

for $h = 0, \ldots, k+1$ and $t \in [0, \widetilde{t}_+]$ with $\widetilde{t}_+$ suitably small provided $y \to R_{k,\delta}(y)$ is a suitable continuous function in the variables δ, y and monotonically increasing in y.

Equation (5.6.91) has the same nature as Eq. (5.6.86), and in the same manner it allows us to show inductively that the Eq. (5.6.49) as well as $||\frac{\partial^j \pi}{\partial x^j}|| \leq \overline{C}$, $j = 0, \ldots, k + 1$, imply

$$\sum_{j=0}^{k+1} ||\frac{\partial^j \pi_{tn}}{\partial x^j}|| \leq (k+1)\frac{R_{k,\delta}((k+1)\overline{C})}{\nu_0/2} + (k+1)\overline{C} \tag{5.6.92}$$

Equation (5.6.92) means that π_∞ is k-times continuously differentiable.

Along similar lines, it is possible to prove the $C^{(k)}$ regularity in the variable α and, jointly, in α and x for $|\alpha|, |x|$ small. By way of estimates of the $(k+1)$-th derivative of π, with respect to α_c of the k-th derivative of π_t with respect to α, ..., and of the first derivative with respect to α of $\frac{\partial^k \pi_t}{\partial x^k}$, this regularity property is proved following the ideas and the techniques of subsections 5.6.M and 5.6.N.

The reader who has been determined enough to reach this point shall not have problems in transforming the above last hints into a proof. We only stress that from what has been said above, it appears that in order to obtain $C^{(k)}$ regularity, one must impose restrictions on δ, a_+, and t_0 which are k dependent. This means that the above proof cannot be used to prove that the attractive manifold depends in a C^∞ way on x and α: actually, it is an open problem to find whether such a smoothness property can be enjoyed by the attractive manifolds under simple extra assumptions (whose necessity is made clear by the example in Observation (1), p.412.) mbe

5.6.13 Exercises

1. Show vague attractivity of $\mathbf{0} = (0,0)$ near $\alpha_c = 0$ for $\dot{x} = \alpha\, x - x^3$, $\dot{z} = -z$, $(x,z) \in \mathcal{R}^2$.

2. In the context of Problem 1, show that the plane $z = 0$ is an attractive manifold in the sense of Proposition 11.

3. Consider the equation in Problem 1 and the surface σ_α built with three pieces with respective parametric equations

$$\begin{cases} z(\gamma) = \overline{z}\, e^{-\gamma} \\ x(\gamma) = \overline{x}(\gamma) = \sqrt{\alpha}\,(1 + \frac{\alpha - \overline{x}^2}{\overline{x}^2} e^{-2\alpha\gamma})^{-\frac{1}{2}} \end{cases}, \qquad \gamma \in [0, +\infty)$$

$$\begin{cases} z(\gamma) = \overline{z}'\, e^{-\gamma} \\ x(\gamma) = \overline{x}'(\gamma) = -\sqrt{\alpha}\,(1 + \frac{\alpha - \overline{x}'^2}{\overline{x}'^2} e^{-2\alpha\gamma})^{-\frac{1}{2}} \end{cases}, \qquad \gamma \in [0, +\infty)$$

$$\begin{cases} z(\gamma) = 0, \\ x(\gamma) = \gamma \end{cases}, \qquad \gamma \in [0, +\infty)$$

Show that σ_α is an attractive manifold $\forall \overline{x}, \overline{x}', \overline{z}, \overline{z}'$ such that $\sqrt{\alpha} \leq \overline{x}, -\overline{x}'$, $\alpha > 0$, in the sense of Proposition 11. (*Hint:* Note that $t \to \overline{x}(t)$ is a solution of $\dot{x} = \alpha x - x^3$ with initial datum $\overline{x}$.)

4. Show that the attractive manifolds in Problem 3 are in $C^{(k)}$, at fixed α, if α is small enough ($2\alpha k < 1$). Show that the equation in Problem 1 admits infinitely many attractive manifolds not C^∞ in x. Meditate on how general this non uniqueness mechanism is.

5. Consider the equation $\dot{x} = \alpha x$, $\dot{z} = -\nu z + x^2$ and determine all the attractive manifolds of the origin for $0 < -a < \nu$. Show that for each $\alpha < 0$ there are infinitely many such manifolds but only one, at most, can be of class C^∞. Find a value of $\alpha < 0$ for which no

attractive manifold is of class C^1. (*Hint:* Note that an attractive manifold must be a union of trajectories of solutions of the differential equation; see also Problem 1. The critical value of α is $\alpha = -\nu/2$.)

6. Using the example of Problem 5 show that the assumption $\mathcal{R}\mathrm{e}\,\lambda_j(\alpha_c) = 0$, $j = 1, \ldots, r$, is essential in Proposition 11. If this assumption is not verified argue that a proposition like Proposition 11 could still hold if the order k of smoothness is restricted as $k < \nu_0/\nu_0'$, at least. See also Problem 7.

7. Prove Proposition 11 for Eq. (5.6.12) when α is near some α_c, $-\nu_0 < \alpha_c < 0$ and $k = 0$. (*Hint:* Write Eq. (5.6.17) as $\dot{x} = \alpha_c x + \chi_\delta(x,z)\Big((a-\alpha_c)x + P(x,z)\Big)$ and proceed as in the proof in §5.6 with the obvious substitution of Eq. (5.6.32), and of the other equations similar to it, with $x = e^{\alpha_c t}x_0 + \int_0^t e^{\alpha_c(t-\tau)}X_\delta(S_\tau^{(\alpha,\delta)}(x_0,\pi(x_0)),\alpha)\,d\tau$, etc.)

8. Show the validity of Proposition 11 in the case in which Eq. (5.6.12) is replaced by the equation ($\mu > 0$)

$$\begin{aligned}\dot{x}_1 =& \alpha x_1 - \mu x_2 + P_1(x_1,x_2,z),\\ \dot{x}_2 =& \mu x_1 + \alpha x_2 + P_2(x_1,x_2,z),\\ \dot{z} =& -\nu_0 z + Q(x_1,x_2,z),\end{aligned}$$

(*Hint:* Write the equation analogous to Eq. (5.6.17) as

$$\begin{aligned}\dot{x}_1 =& -\mu x_2 + \chi_\delta(x_1,x_2,z)(\alpha x_1 + P_1(x_1,x_2,z)),\\ \dot{x}_2 =& \mu x_1 + \chi_\delta(x_1,x_2,z)(\alpha x_2 + P_2(x_1,x_2,z)),\\ \dot{z} =& -\nu_0 z + \chi_\delta(x_1,x_2,z)\,Q(x_1,x_2,z),\end{aligned}$$

with analogous notations. Then proceed exactly as in the proof in §5.6, substituting Eq. (5.6.32), and the other equations similar to it, with

$$\mathbf{x} = W(t)\mathbf{x}_0 + \int_0^t W(t-\tau)\mathbf{X}_\delta(S_\tau^{(\alpha,\delta)}(x_0,\pi(x_0)),\alpha)\,d\tau,$$

where $W(t) = \begin{pmatrix}\cos t & -\sin t\\ \sin t & \cos t\end{pmatrix}$ is the Wronskian matrix; see, also, problems for §2.5, etc.)

9. Using the same ideas as in Problems 7 and 8, study Proposition 11 in the general case, i.e., for an equation of the form of Eq. (5.5.10).

10. If $\mathbf{x}_0$ is not supposed to be vaguely attractive, recognize that the proof of Proposition 11 can be interpreted as showing the existence of a surface σ_α defined as in Eq. (5.6.4), verifying Eq. (5.6.11) and
(ii') If $\mathbf{w} \in \Gamma(\delta_0) \cap \sigma_\alpha$ and $S_t^{(\alpha)}\mathbf{w} \in \Gamma(\frac{1}{2}\delta)$, $\forall\, \tau \in [0,t]$, then $S_t^{(\alpha)}\mathbf{w} \in \sigma_\alpha$ ("local invariance").
(iii') If $S_t\mathbf{w} \in \Gamma(\delta_0) \cap \sigma_\alpha$, $\forall\, t \geq 0$, then $d(S_t^{(\alpha)}\mathbf{w},\sigma_\alpha) \xrightarrow[t\to+\infty]{} 0$ exponentially fast, i.e. the statements of hold as long as the point stays inside $\Gamma(\frac{1}{2}\delta)$. (*Hint:* Vague attractivity is used only to reduce the proof to a theory of (5.6.17), (5.6.18). So just start from them.)

11. Consider the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and suppose that $\mathbf{x}_0 = 0$ is a stationary solution for it. Let L be the stability matrix of $\mathbf{x}_0$ and suppose that L has $(d-r)$ eigenvalues with negative real parts and r with zero real parts. Without imposing the vague attractivity of 0, interpret the proof of Proposition 11 with $\alpha = 0$ as showing that given $k > 0, C > 0$, there exists δ and δ_0, $\delta > \delta_0$, and a surface σ of dimension r and described by $(d-r)$ functions $(\varphi^{(r+1)},\ldots,\varphi^{(d)})$ of r variables $x_1,\ldots,x_r$, with $|x_i| < \frac{1}{2}\delta$ and verifying Eq. (5.6.11) as well

as the local invariance and attractivity properties of the preceding problem ("theorem of the central manifold").

12. Consider the equation $\dot{x} = \lambda x + P(x,z)$, $\dot{z} = -\nu z + Q(x,z)$ with $\lambda, \nu > 0$ and $P, Q \in C^\infty(\mathcal{R}^2)$ with a second order zero at the origin. Show that

$$S_t(x,z) = (e^{\lambda t}x + tD(x,z,t), ze^{-\nu t} + tE(x,z,t))$$

with D and E of class C^∞ and having a zero of second order at $(0,0)$ in the variables (x,z).

13. Use Problem 12 to show that, in the same context and for all small δ, if π is a $C^{(1)}$ function on $[-\delta, \delta]$ such that

$$|\pi(x)| \le \delta, \quad |\frac{d\pi(x)}{dx}| \le \sqrt{\delta} \tag{$*$}$$

and if $\sigma(\pi)$ denotes the curve $z = \pi(x)$, $x \in [-\delta, \delta]$, then $S_t\sigma(\pi)$ is such that $S_t\sigma(\pi) \cap \Gamma(\delta) = \sigma(\pi_t)$. and π_t, verifies Eq. (*). (*Hint:* Use the ideas of the proof of Proposition 11.)

14. In the context of Problems 12 and 13, show that

$$||\pi_{(n+1)\bar{t}} - \pi_{n\bar{t}}|| \le \xi \, ||\pi_{n\bar{t}} - \pi_{(n-1)\bar{t}}||, \qquad ||\pi_{\bar{t}} - \pi'_{\bar{t}}|| \le \xi \, ||\pi - \pi'||$$

with $\xi < 1$ (if $||\cdot||$ denotes the maximum of a function) provided δ_0 is small enough. Deduce the consequent existence in $\Gamma(\delta)$ of a surface locally invariant for S_t and tangent to the x axis at the origin and such that $S_{-t}\mathbf{w} \xrightarrow[t\to+\infty]{} \mathbf{0}$ exponentially fast in the sense $-\lambda = \lim_{t\to+\infty} \frac{1}{t} \log |S_{-t}\mathbf{w}|$ for all nonzero $\mathbf{w}$ on the surface. Denote this surface by σ_i: it is called the "unstable manifold" through $\mathbf{0}$.

15. In the context of Problem 12, show the existence in $\Gamma(\delta)$ of a surface σ_s locally invariant for S_t, tangent to the z axis, and such that $\forall\, \mathbf{w} \ne \mathbf{0}$, $\mathbf{w} \in \sigma_s$ it is $-\nu = \lim_{t\to+\infty} \frac{1}{t} \log |S_t\mathbf{w}|$ ("stable manifold through $\mathbf{0}$").

16. Study the generalization of the result of Problems 10-13 to a general equation in $\mathcal{R}^d$, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ and a stability matrix L whose eigenvalues are pairwise distinct and such that none among them has a zero real part, although some of them have a positive real part and others have a negative real part ("hyperbolic unstable point") ("existence of stable and unstable manifolds at a hyperbolic fixed point").

17. Consider the equation $\dot{x} = x + \frac{x^2}{\alpha} + \frac{z^2}{\beta}$, $\dot{z} = -z + \frac{x^2}{\gamma} + \frac{z^2}{\delta}$, $\alpha = \delta = 1$, $\beta = -\gamma = 2$, and compute the second derivative at the origin of the function π , defining (via $z = \pi_s(z)$) the stable manifold of $\mathbf{0}$. (*Hint:* Write $x = Az^2 + Bz^3 + \ldots$ and insert this expression in the first equation. One finds $A = \beta^{-1}$.)

18. Find some extensions to Problems 14 and 15 to equations in $\mathcal{R}^d$ and study them.

19. In the context of Proposition 11, show that if iσ_α is regarded as an attractor for the neighborhood U used in the proof [see Eq. (5.6.15)], and if $\widetilde{\sigma}_\alpha = \cap_{t>0} S_t^{(\alpha)} \sigma_\alpha$ then the function in the left-hand side of Eq. (5.3.21), p.380, with $A = \widetilde{\sigma}_\alpha$ can be estimated by an exponentially decreasing function of t as $t \to +\infty$, i.e., $\widetilde{\sigma}_\alpha$ is a normal attractor for U by Proposition 5, §5.3, p.379. (*Hint:* Examine the text of Proposition 11 and the discussion around Eq. (5.6.15).)

5.7 An Application: Bifurcations of the Vaguely Attractive Stationary Points into Periodic Orbits. The Hopf Theorem

After the considerations of §5.5 Proposition 10, p.405, the theory of §5.3 can be immediately applied to Eq. (5.1.19). Fixed k, $k \geq 2$, there is $B > 0$ and a cubic neighborhood $\Gamma(\delta)$ centered at $\widehat{\boldsymbol{\omega}}$ with side 2δ, and a family σ_α of $C^{(k)}$ surfaces in $\Gamma(\delta)$ with equations

$$\omega_3 = \widehat{\omega}_3 + \varphi_\alpha(\omega_1, \omega_2), \tag{5.7.1}$$

defined for $|\omega_1|, |\omega_2| \leq \frac{1}{2}\delta$ and α close to α_c, $\alpha \in (\alpha_c - a_+, \alpha_c + a_+)$, and

$$|\varphi_\alpha(\omega_1, \omega_2)| \leq B\,(\omega_1^2 + \omega_2^2) \tag{5.7.2}$$

$\varphi_\alpha \in C^{(k)}([-\frac{1}{2}\delta, \frac{1}{2}\delta]^2 \times (\alpha_c - a_+, \alpha_c + a_+))$, and for every α close to α_c the surface σ_α is invariant in the sense of Eq. (5.6.9) and attractive for all the points of $\Gamma(\delta)$ in the sense of Eq. (5.6.10), with exponential strength.

It will be shown that if $(\alpha - \alpha_c) > 0$ is sufficiently small, there is on σ_α a minimal attractor A_α consisting of a periodic orbit with a period approximately $2\pi/\widehat{\omega}_3$ and attracting the points on $\sigma_\alpha/\{\widehat{\boldsymbol{\omega}}\}$ with exponential strength.

Essentially, by using Proposition 5, §5.3, it will then follow that, in the situation of the preceding sentence, $A_\alpha \cup \{\widehat{\boldsymbol{\omega}}\}$ is an attractor for which the basin $\Gamma(\delta)$ is normal and $\forall\, \boldsymbol{\omega} \in \Gamma(\delta)$, $\exists\, \pi(\boldsymbol{\omega}) \in A_\alpha \cup \{\widehat{\boldsymbol{\omega}}\}$ such that

$$|S_t^{(\alpha)}(\boldsymbol{\omega}) - S_t^{(\alpha)}(\pi(\boldsymbol{\omega}))| \xrightarrow[t \to +\infty]{} 0 \tag{5.7.3}$$

exponentially fast. This statement "completes" the analysis of the asymptotic behavior of the motions of Eq. (5.1.19) with initial datum $\boldsymbol{\omega}$ close enough to $\widehat{\boldsymbol{\omega}}$ and with a given α slightly above α_c.[15]

To see which is the real motion of the gyroscope corresponding to this asymptotically periodic motion of its angular velocity, it would still be necessary to integrate the "geometric" differential equations connecting the Euler angles with the angular velocity, see Eqs. (5.2.9)-(5.2.11). We shall not discuss this last point.

The preceding statements follow, as a special case, from the following general "Hopf bifurcation theorem" and from the observations to it.

12 Proposition. *Consider a differential equation* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ *in* $\mathcal{R}^2$*, parameterized by* $\alpha \in (-a, a)$ *and having the origin* $\mathbf{0}$ *as a vaguely attractive stationary solution near* $\alpha_c = 0$*. Suppose that the stability matrix of the origin, denoted* $L(\alpha)$*, has eigenvalues* $\lambda(\alpha) = \alpha + i\mu(\alpha), \overline{\lambda(\alpha)} = \alpha - i\mu(\alpha)$, $\overline{\mu} \stackrel{def}{=} \mu(0) \neq 0$*. Also suppose that the equation is already put in normal form with respect to*

[15] An even more complete picture, distinguishing the points attracted by A_α from those attracted by $\widehat{\boldsymbol{\omega}}$ can be obtained by using the results of Problems 12-19 of §5.6. The outcome would be the one described just before Proposition 11, p.411.

$\lambda, \overline{\lambda}$ *(see Definition 6, p.392, §5.5; this can always be achieved via a change of coordinates, by Proposition 7, p.393, §5.5):*

$$\begin{aligned} \dot{x} =& \alpha x - \mu(\alpha) y + P(x,y,\alpha) \\ \dot{y} =& \mu(\alpha) x + \alpha y + Q(x,y,\alpha) \end{aligned} \tag{5.7.4}$$

with $P, Q \in C^{(k)}(\mathcal{R}^2 \times (-a,a))$, *$k$ being a large enough integer, and with* P, Q *having a third-order zero in* $x = y = 0$, $\forall\, \alpha \in (-a,a)$.
Finally suppose that the origin is vaguely attractive because the vague attractivity indicator $\overline{\gamma}_{\alpha_c}$ *is negative. Recall that* $\overline{\gamma}_{\alpha_c}$ *is defined as the average value over* θ *of* $\gamma_\alpha(\theta)$ *with*

$$\gamma_\alpha(\theta) = \lim_{\varrho \to 0} \frac{x\, P(x,y,\alpha) + y\, Q(x,y,\alpha)}{(x^2+y^2)^2} \tag{5.7.5}$$

if (ϱ, θ) *are the polar coordinates of* (x,y)*, see (5.5.25).*
Then if $\alpha > 0$ *is sufficiently small, there is a periodic solution to Eq. (5.7.4) which is an attractor attracting all the points in a small neighborhood of* $\mathbf{0}$*, with the exception of* $\mathbf{0}$ *itself, with exponential strength.*
The period T_α *of this motion is such that* $\lim_{\alpha \to \alpha_c} T_\alpha = \frac{2\pi}{\mu(0)}$.

Observations.
(1) The requirement on k to be large enough is imposed to guarantee the possibility of further reducing the complexity of Eq. (5.7.4) by changing coordinates so that the function $\gamma_a(\theta)$ in Eq. (5.7.5) becomes θ independent (i.e. $\gamma_\alpha(\theta) \equiv \overline{\gamma}_\alpha$) in the new polar coordinates and, at the same time, so that in the new coordinates the functions

$$\begin{aligned} r(x,y,\alpha) =& x\, P(x,y,\alpha) + y\, Q(x,y,\alpha) - \overline{\gamma}_\alpha (x^2+y^2)^2, \\ s(x,y,\alpha) =& x\, Q(x,y,\alpha) - y\, P(x,y,\alpha) \end{aligned} \tag{5.7.6}$$

are infinitesimal of fifth order at $x = y = 0$, uniformly in $\alpha \in (-a,a)$, and also have gradients in x, y infinitesimal of the fourth order [a property used below in Eqs. (5.7.17) and (5.7.18)].[16] See Observation (8) for more details.
(2) In the application to Eq. (5.1.19), Eq. (5.7.4) is

$$\begin{aligned} \dot{\omega}_1 =& (\alpha - \alpha_c)\omega_1 - \widehat{\omega}_3\, \omega_2 + P(\omega_1, \omega_2, \alpha), \\ \dot{\omega}_2 =& \widehat{\omega}_3\, \omega_1 + (\alpha - \alpha_c)\omega_2 + Q(\omega_1, \omega_2, \alpha), \end{aligned} \tag{5.7.7}$$

where $P(\omega_1, \omega_2, \alpha), Q(\omega_1, \omega_2, \alpha)$ are respectively

$$\begin{aligned} &- \omega_1(\lambda_2' \omega_1^2 + \lambda_2'' \omega_2^2 + \lambda_2''' 2\widehat{\omega}_3 \varphi_\alpha(\omega_1, \omega_2) + \lambda_2''' \varphi_\alpha(\omega_2, \omega_2)^2) - \omega_2 \varphi_\alpha(\omega_1, \omega_2), \\ &- \omega_2(\lambda_2' \omega_1^2 + \lambda_2'' \omega_2^2 + \lambda_2''' 2\widehat{\omega}_3 \varphi_\alpha(\omega_1, \omega_2) + \lambda_2''' \varphi_\alpha(\omega_2, \omega_2)^2) + \omega_1 \varphi_\alpha(\omega_1, \omega_2), \end{aligned} \tag{5.7.8}$$

[16] $k \geq 5$ will suffice, see Observation (8).

φ_α being the function defining the attractive manifold, and it is of class $C^{(k)}$, k chosen (once and for all) as large as desired. Hence,

$$\gamma_\alpha(\theta) = \lim_{\omega_1,\omega_2 \to 0} -\frac{\lambda_2'\omega_1^2 + \lambda_2''\omega_2^2 + \lambda_2''' 2\widehat{\omega}_3\varphi_\alpha(\omega_1,\omega_2)}{(\omega_0 1^2 + \omega_2^2)} \tag{5.7.9}$$

and to evaluate $\overline{\gamma}_{\alpha_c}$ one does not need to know explicitly φ_α. One can proceed as in the proof of Proposition 10, p.405, setting $r \equiv \varphi_\alpha$; by the same calculation one finds

$$\overline{\gamma}_\alpha = -\frac{(\lambda_2' + \lambda_2'''\widehat{\omega}_3^2)}{2(\lambda_i + 3\lambda_2'''\widehat{\omega}_3^2)} < 0 \tag{5.7.10}$$

Hence, Eq. (5.1.19) has a periodic attractive solution for $\alpha > \alpha_c$ and $(\alpha - \alpha_c)$ small.
(3) As already noted, the assumption that the equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha)$ has normal form with respect to $\lambda(\alpha), \overline{\lambda(\alpha)})$ is not really restrictive if $\mu(0) \neq 0$, by Proposition 7, §5.5, p. 393.
The assumption $\mathcal{R}\mathrm{e}\,\lambda(\alpha) \equiv \alpha$ is also not too restrictive: if $\frac{d\mathcal{R}\mathrm{e}\,\lambda(\alpha)}{d\alpha} \neq 0$ we can rename $\pm\mathcal{R}\mathrm{e}\,\lambda(\alpha)$ with the name α and fall within the assumptions of the theorem. However, pathologies can appear if $\mathcal{R}\mathrm{e}\,\lambda(\alpha)$ has a vanishing derivative at α_c.
(4) The theorem has been formulated in class $C^{(k)}$ rather than in class C^∞ because it is usually applied in connection with the attractive manifold theorem, Proposition 11, p.411 [as, for instance, in Observation (2)], in which case one cannot take $k = +\infty$, in general.
(5) It is important to stress the rather general situation that the above theorem can cover, if combined with the attractive manifold theorem of §5.6, and with the normal-form theorem (Proposition 7, p.393, §5.5) when the loss of stability takes place in two non real conjugate directions. One just has to perform the changes of variables (possible if $\frac{d\mathcal{R}\mathrm{e}\,\lambda(\alpha)}{d\alpha} \neq 0$, $\mu(0) \neq 0$) casting the first two equations, among the d equations of the transformed system, into normal form with respect to the two eigenvalues $\lambda(\alpha), \overline{\lambda(\alpha)}$ "responsible for the loss of stability", as

$$\begin{aligned} \dot{x}_1 =& \alpha x_1 - \mu(\alpha) x_2 + \widetilde{P}(x_1, x_2, \mathbf{y}, \alpha), \\ \dot{x}_2 =& \mu(\alpha) x_1 + \alpha x_2 + \widetilde{Q}(x_1, x_2, \mathbf{y}, \alpha), \end{aligned} \tag{5.7.11}$$

where $\mathbf{y}$ denotes the remaining $(d-2)$ unknowns of the differential equation.

Then one considers the differential equation in $\mathcal{R}^2$ of the form of Eq. (5.7.4) with $P(x_1, x_2, \alpha) = \widetilde{P}(x_1, x_2, \mathbf{0}, \alpha), Q(x_1, x_2, \alpha) = \widetilde{Q}(x_1, x_2, \mathbf{0}, \alpha)$. If this equation verifies the assumptions of Proposition 12, we can infer that the original equation has an attractive periodic orbit for α slightly above α_c.
The proof of this simple criterion is obtained by the obvious extension to $\mathcal{R}^d$ of the discussion in Observation (2) (write $\mathbf{y} = \boldsymbol{\varphi}_\alpha(x_1, x_2)$ and use the fact that $\boldsymbol{\varphi}_\alpha$ vanishes to second order.

(6) The above theorem has a natural analogue in one dimension. Consider the equation in $\mathcal{R}$:

$$\dot{x} = \alpha x + p(x, \alpha), \tag{5.7.12}$$

where $p \in C^{(k)}(\mathcal{R}^2)$, k large enough, and p has a third-order zero in $x = 0$, $\forall\, \alpha \in (-a, a)$, and

$$c(\alpha) = \lim_{x \to 0} \frac{x\, p(x, \alpha)}{x^4} < 0 \tag{5.7.13}$$

with $c(\alpha) < 0$ and continuous near $\alpha = 0$. If k is large, by the implicit function theorem, Eq. (5.7.12) has two stationary solutions, for $\alpha > 0$ and small ($x \simeq \pm\frac{\alpha}{c(\alpha)}$.

At such points, the stability "matrix" is $-2\alpha < 0$ and, therefore, the two points are attractors with exponential strength for the points in their vicinity.

This observation is sometimes useful in treating cases analogous to the ones discussed in Observation (5), when the stationary solution loses stability because only one real eigenvalue crosses the imaginary axis, as a grows through a critical value α_c leaving the stationary solution vaguely attractive.

However, it should be stressed that this is a rather rare possibility since it is generally impossible to put a one-dimensional equation into normal form, see observation (3), p.397. The existence of normal form can be expected only in systems with "some symmetry" (like $x \longleftrightarrow -x$ odd symmetry of $p(x, \alpha)$).

Note also that if Eq. (5.7.12) has the property of Eq. (5.7.13), then a small perturbation of it, like

$$\dot{x} = \alpha x + p(x, \alpha) + \varepsilon x^2, \tag{5.7.14}$$

can change the vague-attractivity character of $x = 0$ for α near 0, no matter how small ε is (exercise). This phenomenon is not possible in equations in which the loss of stability takes place in two complex non real directions (essentially just because of the existence of normal forms).

(7) The mechanism of generation of a periodic orbit out of a fixed point when a grows through α_c described in Proposition 12, is called a "Hopf bifurcation". The solution $\mathbf{x}_0$ loses stability in two complex directions at $\alpha = \alpha_c$ and, if it stays vaguely attractive in the sense of Eq. (5.7.5), it is surrounded by a periodic attractive motion taking place on a curve whose diameter, as we shall see, grows as $\sqrt{\alpha - \alpha_c}$ for $\alpha - \alpha_c > 0$ and small.

(8) As shown in the proof of Proposition 8, p.396, it is always possible to change smoothly coordinates so as to put Eq. (5.7.4) into a form such that $\overline{\gamma}_\alpha(\theta)$ is θ independent: $\gamma_\alpha(\theta) \equiv \overline{\gamma}_\alpha$, $\forall\, \alpha \in (-a, a)$, i.e.,

$$\begin{aligned} \dot{x} &= \alpha x - \mu(\alpha) y + \overline{\gamma}_\alpha x(x^2 + y^2) + \widetilde{P}(x, y, \alpha), \\ \dot{y} &= \mu(\alpha) x + \alpha y + \overline{\gamma}_\alpha y(x^2 + y^2) + \widetilde{Q}(x, y, \alpha), \end{aligned} \tag{5.7.15}$$

with $\overline{\gamma}_0 < 0$ and with P, Q infinitesimal of fourth order at $x = y = 0$, uniformly in $\alpha \in (-a, a)$ (possibly reducing the value of α); see the change of variables of Eq. (5.5.39) changing Eq. (5.5.38) (i.e., essentially, Eq. (5.7.4) written in complex form) into Eq. (5.5.43) (ie., (5.5.37)). By Eqs. (5.5.42) and (5.5.38), it one realizes that the needed change of coordinates involves the third-order Taylor coefficients of P and Q at $x = y = 0, \alpha = \alpha_c$, with respect to the variables x, y and it turns out to be of class C^∞ in the variables x, y near $x = y = 0$ and α small (but, in general, only of class $C^{(k-3)}$ in α).
If $k > 5$, the functions P, Q in Eq. (5.7.15) have fifth-order derivatives with respect to x, y continuous in x, y, a near $(0, 0, 0)$ and also have a fourth-order zero in x, y at $x = y = 0$, $\forall\, \alpha \in (-a, a)$, if a is small.
Furthermore, the functions r, s of Eq. (5.7.6) are now equal to

$$\begin{aligned} r(x, y, \alpha) =& x\, \widetilde{P}(x, y, \alpha) + y\, \widetilde{Q}(x, y, \alpha), \\ s(x, y, \alpha) =& x\, \widetilde{Q}(x, y, \alpha) - y\, \widetilde{P}(x, y, \alpha), \end{aligned} \tag{5.7.16}$$

by Eq. (5.7.15), and their derivatives in x, y are continuous in x, y, α near $(0, 0, 0)$ and have a fourth-order zero at $x = y = 0$, $\forall\, \alpha \in (-a, a)$.
Hence, to fix the ideas, we shall suppose that "k large enough" means $k > 5$. However, this is not optimal, and one can improve the value of the degree of regularity in x, y, α necessary for P, Q so that a proposition like Proposition 12 will hold in general. To obtain fine results, one should distinguish the regularity imposed on the α variable and that on the x, y variables.

Proof. By observation (8), if $k > 5$, it suffices to treat Eq. (5.7.15) with $\widetilde{P}, \widetilde{Q}, \partial r, \partial s$ [see Eqs. (5.7.15) and (5.7.16)] being fourth-order infinitesimals in x, y for $x = y = 0$, uniformly in $\alpha \in (-a, a)$ (here ∂ denotes the gradient with respect to the x, y variables).

Let $\overline{\gamma} \equiv -\overline{\gamma}_0, \overline{\mu} \equiv \mu(0)$. By the infinitesimality properties of $\widetilde{P}, \widetilde{Q}$, it is possible to find $\overline{\varrho} > 0, 0 < \overline{a} < a$, such that, for all $(x, y) \in C(\overline{\varrho})/\{\mathbf{0}\}$, with $C(\overline{\varrho}) \overset{def}{=} \{x, y \,|\, (x, y) \in \mathcal{R}^2,\ \sqrt{x^2 + y^2} \le \overline{\varrho}\}$, and for all $\alpha \in (-\overline{a}, \overline{a})$

$$\begin{aligned} &\alpha - \frac{1}{8}\overline{\gamma}\overline{\varrho}^2 < 0, \qquad \frac{2}{3}\overline{\mu} < \mu(\alpha) < \frac{3}{2}\overline{\mu}, \\ &-2\overline{\gamma} < \overline{\gamma}_\alpha + \frac{r(x, y, \alpha)}{(x^2 + y^2)^2} < -\frac{\overline{\gamma}}{2}, \qquad \frac{|s(x, y, \alpha)|}{(x^2 + y^2)^2} < \frac{\overline{\mu}}{2} \end{aligned} \tag{5.7.17}$$

having supposed, for definiteness, that $\overline{\mu} > 0$. Call $C(\varrho', \varrho'') \overset{def}{=} \{$annulus with radii $\varrho' < \varrho'' = \{x, y \,|\, (x, y) \in \mathcal{R}^2,\ \varrho' < \sqrt{x^2 + y^2} < \varrho''\}$

We now check that Eq. (5.7.17) implies that the disk $C(\overline{\varrho})$ is $S_t^{(\alpha)}$ invariant and that there is also an invariant annulus $C(\varrho'_\alpha, \varrho''_\alpha) \subset C(\overline{\varrho})$ with $0 < \varrho'_\alpha, \varrho''_\alpha < \overline{\varrho}$ which is an attractor for the points in $C(\overline{\varrho})/\{\mathbf{0}\}$, for all $\alpha \in (-\overline{a}, \overline{a})$.

In fact, multiply the first of Eqs. (5.7.15) by x, the second by y, and add the results:

$$\begin{aligned}\frac{d}{dt}\frac{x^2+y^2}{2} &= \big(\alpha+\overline{\gamma}_\alpha(x^2+y^2)+\frac{r(x,y,\alpha)}{(x^2+y^2)^2}\big)(x^2+y^2)\\ &= \begin{cases} < (\alpha-\overline{\gamma}\frac{x^2+y^2}{2})(x^2+y^2)\\ > (\alpha-2\overline{\gamma}(x^2+y^2))(x^2+y^2)\end{cases}\end{aligned} \tag{5.7.18}$$

which shows [see the first of Eqs. (5.7.17)] that the intermediate term in Eq. (5.7.18) is negative on $\partial C(\overline{\varrho})$. This means that $C(\overline{\varrho})$ is $S_t^{(\alpha)}$ invariant, $\forall\, t \geq 0$, $\forall\, \alpha \in (-\overline{a}, \overline{a})$. Let

$$\overline{\varrho}'_\alpha = \sqrt{\frac{\alpha}{2\overline{\gamma}}}, \qquad \overline{\varrho}''_\alpha = \sqrt{\frac{2\alpha}{\overline{\gamma}}} \tag{5.7.19}$$

and note that the inequalities in Eq. (5.7.18) show that the intermediate term in Eq. (5.7.18) is positive on $\partial C(\overline{\varrho}'_\alpha)$ and negative on $\partial C(\overline{\varrho}''_\alpha)$; hence, the annulus $C(\overline{\varrho}'_\alpha, \overline{\varrho}''_\alpha)$ is $S_t^{(\alpha)}$ invariant, if α is small so that $\overline{\varrho}''_\alpha < \overline{\varrho}$.

Equations (5.7.17) and (5.7.18) also show that if $\varrho'_\alpha = \frac{1}{2}\overline{\varrho}'_\alpha$, $\varrho''_\alpha = 2\overline{\varrho}''_\alpha < \overline{\varrho}$ the annulus $C(\varrho'_\alpha, \varrho''_\alpha)$ is also invariant and enjoys the property that any initial datum chosen in $C(\overline{\varrho})/\{0\}$ evolves, entering into $C(\varrho'_\alpha, \varrho''_\alpha)$ in a finite time (see Fig. 5.7), $\forall\, \alpha \in (-\overline{a}, \overline{a})$.

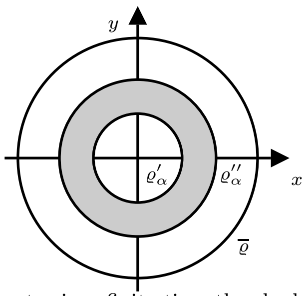


Figure 5.7: Initial data in $C(\overline{\varrho})$ enter in a finite time the shaded annulus $C(\varrho'_\alpha, \varrho''_\alpha)$.

In fact, if $\overline{\varrho} > \sqrt{x^2+y^2} > \varrho''_\alpha$, the first inequality in the right-hand side of Eq. (5.7.18) shows that the intermediate term of Eq. (5.7.18) is $\leq -\frac{6\alpha^2}{\overline{\gamma}^2}$, so that the "entrance time" in $C(\varrho'_\alpha, \varrho''_\alpha)$ is finite and can be estimated by $\tau = \frac{\overline{\varrho}^2 - \overline{\varrho}''^2_\alpha}{12\alpha^2}$.

If $0 < \widetilde{\varrho} = \sqrt{x^2+y^2} < \varrho'_\alpha$ the intermediate term of Eq. (5.6.18) is not less than $m = \min \varrho'_\alpha \geq \varrho \geq \widetilde{\varrho}(\alpha\varrho^2 - 2\overline{\gamma}\varrho^4) > 0$ by the second inequality in the right-hand side of Eq. (5.6.18). Hence, the entrance time can now be estimated by $\tau = ({\varrho'_\alpha}^2 - \widetilde{\varrho}^2)/2m$.

This means that every datum close to the origin moves away from the origin until it enters the annulus $C(\varrho'_\alpha, \varrho''_\alpha)$ in a finite time, while every datum close to $\partial C(\overline{\varrho})$ moves towards the origin until it enters the annulus $C(\varrho'_\alpha, \varrho''_\alpha)$ in a finite time. These motions are spiraling motions, as we now show.

To see that the motions starting in $C(\overline{\varrho})/\{\mathbf{0}\}$ are "spiraling motions", it suffices to study them in polar coordinates.

If $S_t^{(\alpha)}(x,y) \equiv (x(t),y(t))$ and if $(\varrho(t),\theta(t))$ are the polar coordinates of $(x(t),y(t)) \in C(\overline{\varrho})/\{0\}$,

$$\begin{aligned} \frac{d\theta}{dt} &= \frac{d}{dt}\mathrm{arctg}\frac{y(t)}{x(t)} = \frac{\dot{y}x - \dot{x}y}{x^2+y^2}, \\ \frac{d\varrho}{dt} &= \frac{d}{dt}\sqrt{x(t)^2+y(t)^2} = \frac{\dot{x}x+\dot{y}y}{\sqrt{x^2+y^2}}, \end{aligned} \tag{5.7.20}$$

Note that if $\varrho(0) > 0, \varrho(0) < \overline{\varrho}$, then $\varrho(t) > 0$ and $\varrho(t) < \overline{\varrho}$ for all $t \geq 0$, because of the above arguments. Hence, Eq. (5.7.15) and the second and fourth inequalities in (5.7.17) imply

$$\dot{\theta} = \mu(\alpha) + \frac{s(x,y,\alpha)}{(x^2+y^2)^2} \;\Rightarrow\; \frac{1}{4}\overline{\mu} < \dot{\theta} < 2\overline{\mu}, \tag{5.7.21}$$

i.e., θ is monotonic in t and diverges as $t \to +\infty$. This just means that the motion spirals if $0 < \varrho(0) < \overline{\varrho}$.

We now check that the spirals associated with the initial data external to $C(\varrho''_\alpha)$, but in $C(\overline{\varrho})$, become asymptotically confused, as $t \to +\infty$, with those associated with data internal to $C(\varrho'_\alpha)$, but different from the origin.

If this happens, the two families of spirals are separated by a periodic orbit which will be an attractor with basin containing $C(\overline{\varrho})/\{0\}$.

To discuss the asymptotic identity of the spirals it is convenient to describe them as geometric objects, thinking of them as parameterized in terms of θ instead of t, which is possible by Eq. (5.7.21).

Let $\theta \to \varrho_1(\theta)$ and $\theta \to \varrho_2(\theta)$ be the equations in polar coordinates of two spirals on which two motions of Eq. (5.7.15) run, starting with initial data $\varrho_1(0) \geq \varrho'_\alpha$, $\theta_1(0) = 0$ and $\varrho_2(0) \leq \varrho''_\alpha$, $\theta_2(0) = 0$ and $\varrho_1(0) < \varrho_2(0)$.

By the uniqueness theorem for the solutions of the differential equations and by the autonomy of Eq. (5.7.15), we see that $\varrho_2(\theta) - \varrho_1(\theta) > 0$, $\forall\,\theta \geq 0$. We show the existence of $R > 0, \varepsilon(\alpha) > 0$ such that for α small enough,

$$\varrho_2(\theta)\varrho_1(\theta) \leq R\,e^{-\varepsilon(\alpha)\theta} \tag{5.7.22}$$

Then the autonomy of Eq. (5.7.15) and Eqs. (5.7.22) and (5.7.21) plus the attractivity properties of $C(\varrho'_\alpha, \varrho''_\alpha)$ will imply that every datum in $C(\overline{\varrho})/\{0\}$ evolves exponentially fast in θ (with rate constant $> \varepsilon(\alpha) > \frac{1}{6}\overline{\mu}$) towards a periodic trajectory of Eq. (5.7.15) which separates geometrically the "outer" spirals (i.e., those originating outside $C(\varrho''_\alpha)$) from the "inner" spirals (i.e., those originating inside $C(\varrho'_\alpha)$).

To prove Eq. (5.7.22), note that Eqs. (5.7.20), (5.7.19), and (5.7.21) imply

$$\frac{d\varrho}{dt} = \varrho\frac{\alpha\,(x^2+y^2) + \overline{\gamma}_\alpha(x^2+y^2)^2 + r(x,y,\alpha)}{\mu(\alpha)\,(x^2+y^2) + s(x,y,\alpha)} \tag{5.7.23}$$

where r, s are infinitesimals of fifth order in x, y at $x = y = 0$, uniformly in $\alpha \in (-\overline{a}, \overline{a})$, while their gradients with respect to x and y have the same property to fourth order.

Equation (5.7.23) will be rewritten as

$$\frac{d \log \varrho}{dt} = \frac{\alpha + \overline{\gamma}_\alpha \varrho^2 + r(x,y,\alpha)\varrho^{-2}}{\mu(\alpha) + s(x,y,\alpha)\varrho^{-2}} \tag{5.7.24}$$

We now wish to show that the right-hand side of Eq. (5.7.24) is monotonic in ϱ for $\varrho \in [\varrho'_\alpha, \varrho''_\alpha]$ at fixed θ and that its ϱ derivative stays away from zero.

To estimate the derivative just compute it. Basically, the possibility of the bound is due to the fact that to the lowest order in ϱ, the right-hand side of Eq. (5.7.24) is $(\alpha + \overline{\gamma}_\alpha \varrho^2)/\mu(\alpha)$ whose ϱ-derivative is $2\overline{\gamma}_\alpha \varrho/\mu(\alpha)$.

So we expect that if $\overline{\varrho}$ is small enough [with $\overline{a}$ chosen correspondingly small so that the first of Eqs. (5.7.17) still holds], the ϱ derivative of the right-hand side of Eq. (5.7.24) can be estimated, $\forall\, \varrho \in [\varrho'_\alpha, \varrho''_\alpha]$ (using the orders of infinitesimality of $r, s, \partial r, \partial s$ neglect the terms in r, s) to be not larger than:

$$-\frac{\overline{\gamma}}{\overline{\mu}}\sqrt{\frac{1}{2\overline{\gamma}}}\sqrt{\alpha} \equiv -\chi\sqrt{\alpha} \tag{5.7.25}$$

A direct calculation of the ϱ derivative of the right-hand side of Eq. (5.7.24) actually proves the above statement, by Eq. (5.7.25).

Then recalling that $\varrho_2(\theta) > \varrho_1(\theta), \forall\, \theta > 0$, and writing Eq. (5.7.24) for ϱ_2 and ϱ_1, and subtracting them, we find, applying the bound on the derivative (5.7.25) (recalling that $\varrho'_\alpha \le \varrho_1(\theta)$):

$$\begin{aligned} \frac{d}{d\theta}\log\frac{\varrho_2(\theta)}{\varrho_1(\theta)} &\le -\chi\sqrt{\alpha}(\varrho_2(\theta) - \varrho_1(\theta)) = -\chi\sqrt{\alpha}\varrho_1(\theta)\Big(\frac{\varrho_2(\theta)}{\varrho_1(\theta)} - 1\Big) \\ &\le -\chi\sqrt{\alpha}\varrho'_\alpha\Big(\frac{\varrho_2(\theta)}{\varrho_1(\theta)} - 1\Big) = -\frac{\alpha}{2\overline{\mu}}\Big(\frac{\varrho_2(\theta)}{\varrho_1(\theta)} - 1\Big) \end{aligned} \tag{5.7.26}$$

which interpreted as a differential inequality for $\frac{\varrho_2(\theta)}{\varrho_1(\theta)}$, yields

$$\Big(1 - \frac{\varrho_1(\theta)}{\varrho_2(\theta)}\Big) \le \Big(1 - \frac{\varrho_1(0)}{\varrho_2(0)}\Big)\, e^{-\frac{\alpha}{2\overline{\mu}}\theta} \tag{5.7.27}$$

by integration, and this completes the proof. mbe

5.7.1 Exercises and Problems

1. The estimate for the coefficient $\varepsilon(\alpha)$ in Eq. (5.7.22) is [see Eq. (5.7.27)], $\varepsilon(\alpha) = \frac{\alpha}{2\overline{\mu}}$. Is it possible to improve it so that the new estimate $\widetilde{\varepsilon}(\alpha)$ has the property that $\widetilde{\varepsilon}(\alpha) \xrightarrow[\alpha \to 0]{} \varepsilon > 0$? If not, find a physical interpretation or a motivation of this fact.

2. Consider the differential equation in $\mathcal{R}^2$ written in complex form as $\dot{z} = \xi(\alpha)z + P(z, \overline{z})$, where $z = x + iy$, $(x,y) \in \mathcal{R}^2$, $\xi(\alpha) = \sigma(\alpha) + i\mu(\alpha)$, and let $\sigma(0) = 0, \mu(0) \ne 0$, $\sigma, \mu \in C^\infty(\mathcal{R}^2)$; suppose P to be a C^∞ function of x, y with a second-order zero at the origin. In the proof of Proposition 8, p.396, it was shown [see the change of variables in Eq. (5.5.39)] that in some new coordinates the equation can be given the form $\dot{z} = \xi(\alpha)z + c_2 z(\alpha) z|z|^2 + O(|z|^4)$,

where $O(|z|^4)$ symbolically denotes a function of x, y, α of class C^∞ and with a fourth-order zero at $z = 0$ for all α near zero. Show that the equation can be given the form:

$$\dot{z} = \xi(\alpha)z + c_2(\alpha)z|z|^2 + +O(|z|^4)$$

with the same meaning of the symbols, after a new change of coordinates. (*Hint:* Again change coordinates as $\zeta = z + \Gamma_4(z, \overline{z})$, where Γ_4 is a homogeneous polynomial in $z, \overline{z}$ of fourth degree, such that the fourth-order terms in the equation cancel, see Eq. (5.5.39)-(5.5.43).)

3. In the context of Problem 2, develop the same ideas to show that, $\forall\, k > 0$, the equation can be put, in a suitable coordinate system, in the form

$$\dot{z} = \xi(\alpha)z + c_2(\alpha)z|z|^2 + c_4 z|z|^4 + \ldots c_{2k} z|z|^{2k} + O(|z|^{2(k+1)})$$

(*Hint:* Use induction.)

4. Show that in Problems 2 and 3, the assumption $\sigma(0) = 0$ is not necessary. Actually, if $\sigma(0) \neq 0$, show that, by the same type of arguments, the equation can be given the form

$$\dot{z} = \xi(\alpha)z + O(|z|^k)$$

for all $k > 0$. (*Hint:* Note that the reason why one could not eliminate $c_2 z|z|^2$ in Problem 2 was that $\lambda(0) + \overline{\lambda(0)} = 0$.)

5. In Problems 2-4, the parameter α does not play a very essential role. Formulate statements of the same type for α-independent equations. (*Hint:* Just set $\alpha = 0$ in Problems 2-4 and determine what can be said.)
For information about the problems related to the iterated composition of coordinate transformations transforming the original equations into a fully linear equation $\dot{z} = \xi z$ when $\mathcal{R}\mathrm{e}\, \xi \neq 0$ and $\mathcal{I}\mathrm{m}\, \xi \neq 0$ (by letting $k \to +\infty$ in Problem 4), see [34].

6. Discuss the bifurcation pattern, as a grows, for the stationary solutions of the equation

$$\begin{aligned}
\dot{\gamma}_1 &= -\,\gamma_1 + 4\gamma_2\gamma_3,\\
\dot{\gamma}_2 &= -\,9\gamma_2 + 3\gamma_1\gamma_3,\\
\dot{\gamma}_1 &= -\,5\gamma_1 - 7\gamma_1\gamma_2 + \alpha,
\end{aligned}$$

7. Same as problem 6 for

$$\begin{aligned}
\dot{\gamma}_1 &= -\,2\gamma_1 + 4\gamma_2\gamma_3 + 4\gamma_4\gamma_5,\\
\dot{\gamma}_2 &= -\,9\gamma_2 + 3\gamma_1\gamma_3,\\
\dot{\gamma}_3 &= -\,5\gamma_3 - 7\gamma_1\gamma_2 + \alpha,\\
\dot{\gamma}_4 &= -\,5\gamma_4 - \gamma_1\gamma_5,\\
\dot{\gamma}_5 &= -\,\gamma_5 - 3\gamma_1\gamma_4.
\end{aligned}$$

assuming (without checking it) that when a stationary solution loses stability in one real direction or in two complex ones, it remains vaguely attractive with a negative vague-attractivity indicator [as defined in Eqs. (5.5.24) and (5.5.25)]. See §5.8 for a more detailed analysis.

8. Find some improvements on the regularity requirements in the variables x, y, and α in Proposition 12, possibly requiring a different order of regularity in x, y, or α.

9. In the context of Proposition 12, suppose that $\overline{\gamma}_{\alpha_c}$, as defined there, is positive. Show that in this case, if $\alpha_c = 0$, there is a repulsive periodic orbit for Eq. (5.7.4) for $\alpha < 0$ small. (*Hint:* Just change t into $-t$ and apply Proposition 12, noting that the change of t into $-t$ changes the notion of attractivity into that of "repulsivity".)

5.8 On the Stability Theory for Periodic Orbits and More Complex Attractors (Introduction)

Nondum matura est.

In this section we devote some attention to what happens, as α increases, to the periodic solution of Eq. (5.1.19) whose existence has been established in §5.6 and §5.7. More generally, one can ask how to establish stability criteria for periodic solutions to differential equations, with uniformly bounded trajectories, of the type:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \alpha) \tag{5.8.1}$$

with $\mathbf{f} \in C^\infty(\mathcal{R}^d \times \mathcal{R})$ or $\mathbf{f} \in C^{(k)}(\mathcal{R}^d \times \mathcal{R})$ with k large enough.

Before examining the evolution of the stability of a periodic orbit of Eq. (5.8.1) when α varies, it is necessary to investigate the notions of stability of a periodic motion of the equation in $\mathcal{R}^d$:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{5.8.2}$$

with $\mathbf{f} \in C^\infty(\mathcal{R}^d)$ or $C^{(k)}(\mathcal{R}^d)$ with k large enough and such that Eq. (5.8.2) has bounded trajectories.

Let $t \to \mathbf{x}(t)$, $t > 0$, be a periodic solution of Eq. (5.8.2) with minimal period $T > 0$. The stability and the attractivity of this solution is conveniently described in terms of the "Poincaré transformation".

7 Definition. *Let $t \to \mathbf{x}(t)$ be a periodic solution of Eq. (5.8.2) with minimal period $T > 0$.*
Let $\boldsymbol{\xi}_0$ be a point on this trajectory, say $\boldsymbol{\xi}_0 = \mathbf{x}(0) \in \mathcal{R}^d$, and let σ be a $(d-1)$-dimensional flat surface element cutting the orbit at the point to so that the orbit is not tangent to σ in $\boldsymbol{\xi}_0$ ("transversal surface element").
It is then possible to define a C^∞ transformation [or a $C^{(k)}$ transformation, if the right-hand side of Eq. (5.8.2) is only of class $C^{(k)}$], on a neighborhood of $\boldsymbol{\xi}_0$ relative to σ and with values on σ itself, by considering a neighborhood U of $\boldsymbol{\xi}_0$ on σ so small that the motion, according to Eq. (5.8.2), of the initial datum $\boldsymbol{\xi} \in U$ comes back to intersect σ for the first time after a time $T_{\boldsymbol{\xi}} \simeq T$ at a point $\Phi_\sigma(\boldsymbol{\xi}) \in \sigma$.
The map of $\sigma \cap U$ into σ associating with $\boldsymbol{\xi} \in \sigma \cap U$ the point $\Phi_\sigma(\boldsymbol{\xi}) \in \sigma$ is called the "Poincaré transformation" relative to the given periodic orbit, to the given surface element, and to the given vicinity U.

It is then possible to formulate the following sufficient stability and attractivity criterion (and instability criterion as well) for a periodic orbit. It is the best illustration of the meaning and of the interest of the Poincaré maps.

13 Proposition. *Let* $t \to \mathbf{x}(t)$, $t \geq 0$, *be a periodic motion for Eq. (5.8.2) with minimal period* $T > 0$.
Let σ *be a transversal surface element to the trajectory in* $\boldsymbol{\xi}_0 = \mathbf{x}(O)$ *and introduce on a Cartesian coordinates* $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{d-1})$ *with origin in* $\boldsymbol{\xi}_0$. *Denote by* $\boldsymbol{\eta}' = \widehat{\Phi}_\sigma(\boldsymbol{\eta})$ *the Poincaré map defined in a suitable neighborhood of* $\boldsymbol{\xi}_0$ *on* σ. *By definition it is* $\widehat{\Phi}_\sigma(\mathbf{0}) = \mathbf{0}$.
Define the "stability" or "Lyapunov" matrix of the periodic orbit, relative to σ *and to the given system of coordinates on it, as*

$$(L_\sigma)_{ij} = \frac{\partial \widehat{\Phi}_\sigma^{(i)}}{\partial \eta_j}(\mathbf{0}), \qquad i, j = 1, \ldots, d-1 \tag{5.8.3}$$

Then the periodic orbit is stable and is an attractor, with exponential strength, for the points close enough to it if all the eigenvalues of the matrix L_σ *have modulus less than* 1.
If at least one among the eigenvalues of Le has modulus larger than 1, the orbit is unstable.

Observations.
(1) This proposition is analogous to Proposition 6, p.382, §5.4, formulated for maps rather than for differential equations (which can, however, be thought of as "infinitesimal maps"). Its proof is left to the reader as an interesting problem [see also Observation (2) below]. To study it, one should first understand the case when $\widehat{\Phi}_\sigma$ is a linear map near $\boldsymbol{\xi}_0$. Proposition 13 bears the name "stability criterion of Lyapunov" for maps.
(2) Proposition 13 is a special case of a slightly different proposition which could be formulated on the stability of stationary points with respect to the action of repeated applications of a map of $\mathcal{R}^d$ into itself.
The fact that $\widehat{\Phi}_\sigma$ is a Poincaré map plays little role in the proof which is, in fact, split into two parts:
(i) show that the origin is an exponentially attracting (or, alternatively, unstable) point for the iterates of $\widehat{\Phi}_\sigma$;
(ii) remark that since $\widehat{\Phi}_\sigma$ is a Poincaré map relative to a periodic orbit for Eq. (5.8.2), (i) implies that the periodic orbit exponentially attracts the points close enough to it (or is, alternatively, unstable).
And (ii) follows trivially from (i), which could be phrased without reference to the Poincaré map but simply for an arbitrary map of a surface into itself (with a fixed point).

Now consider Eq. (5.8.1) and assume that, $\forall\, \alpha \in (\alpha', \alpha'') \stackrel{def}{=} J$, this equation admits among its solutions a periodic motion $t \to \mathbf{x}_\alpha(t)$, $t \geq 0$, with minimal period $T_\alpha > 0$ and such that the function $(\alpha, t) \to \mathbf{x}_\alpha(t)$ is a $C^{(k)}$

function on $J \times [0, +\infty)$, if $C^{(k)}$ is the regularity class in the right-hand side of Eq. (5.8.1).

It will then be possible to consider, $\forall\, \alpha \in (\alpha', \alpha'')$, the stability matrix $L_\sigma(\alpha)$, see Eq. (5.8.3), relative to a surface element σ which, if $J = (\alpha', \alpha'')$ is a small enough interval, can be supposed to be α independent.

We can choose the Cartesian coordinate system on σ for each α, with the origin at the point $\boldsymbol{\xi}_\alpha$ at the intersection of σ and the trajectory, and smoothly varying with α so that the Poincaré maps $\widehat{\Phi}_{\sigma,\alpha}(\boldsymbol{\eta})$ are defined for $\boldsymbol{\eta} \in U$, where U is a small enough neighborhood of the origin, and $\widehat{\Phi}_{\sigma,\alpha}(\boldsymbol{\eta})$ is of class $C^{(k)}$ on $U \times (\alpha', \alpha'')$ in the variables $(\boldsymbol{\eta}, \alpha)$ and

$$\widehat{\Phi}_{\sigma,\alpha}(\mathbf{0}) = \mathbf{0}, \qquad \forall\, \alpha \in J. \tag{5.8.4}$$

We can and shall suppose that $\widehat{\Phi}_{\sigma,\alpha}$ is extended arbitrarily to a map of $\mathcal{R}^{d-1}$ into itself, having the same regularity class $C^{(k)}$ (to define this extension it might be first necessary to reduce slightly the size of U).

In analogy with the definitions of stability, attractivity, etc. relative to the solution flows associated with differential equations, we can introduce analogous notions for a single transformation Φ of $\mathcal{R}^d$, or of an open subset of $\mathcal{R}^d$, into itself. What was formerly the family $(S_t)_{t\geq 0}$ of maps associated with the solution of the differential equation now becomes the family $(\Phi^n)_{n\in\mathcal{Z}_+}$ of the iterations of Φ, i.e., one can think of Φ as an "evolution" on $\mathcal{R}^d$ observed at integer times.

We do not repeat the obvious process of setting up the notions of stability, attractivity, vague attractivity, etc. for the iterations of a map Φ, and we just mention that once such definitions are posed in an obvious manner (taking into account the analogous definitions associated with the differential equations), the following proposition on the existence of an attractive manifold and on the Hopf bifurcations holds.

14 Proposition. *(i) Consider Eq. (5.8.1) with* $\mathbf{f} \in C^{(k+1)}$, $k \geq 1$, *and suppose that the equation admits a family of periodic orbits verifying the properties illustrated in the above text, following the observations to Proposition 13.*

Suppose that for $\alpha \in J \stackrel{def}{=} (\alpha', \alpha'')$, *the stability matrix* $L_\sigma(\alpha)$ *has the eigenvalues* $\lambda_{s+1}(\alpha), \ldots, \lambda_{d-1}(\alpha)$ *with modulus less or equal to* $\nu < 1$ *and that, for some* $\nu' \in (\nu, 1)$, *the other eigenvalues* $\lambda_1(\alpha), \ldots, \lambda_s(\alpha)$ *have modulus larger or equal to* ν'. *Also suppose that the plane generated by the eigenvectors of* $L_\sigma(\alpha)$ *associated with the eigenvalues* $\lambda_1(\alpha), \ldots, \lambda_s(\alpha)$ *coincides with the plane* $\eta_{s+1} = \ldots = \eta_{d-1} = 0$.

If the origin is vaguely attractive for the maps $\widehat{\Phi}_{\sigma,\alpha}$ *near* $\alpha_c \in J$, *and if* $\lambda_j(\alpha_c)| = 1$, $j = 1, \ldots, s$, *there exist* $\varepsilon > 0, \delta, \delta_0 > 0, \delta_0 < \delta$ *and* $d-1-s$ *functions* $\varphi^{(s+1)}, \ldots, \varphi^{(d-1)}$ *defined in the neighborhood*[17] $\Gamma_s(\frac{\delta}{2}) \times (\alpha_c - \varepsilon, \alpha_c + \varepsilon)$ *and there of class* $C^{(k)}$ *such that the equations*

[17] As usual, $\Gamma_s(\delta) = \{\mathbf{x}|\mathbf{x} \in \mathcal{R}^s,\ |x_i| < \delta,\ i = 1, \ldots, s.$

$$\eta_{s+j} = \varphi^{(s+j)}(\eta_1, \ldots, \eta_{d-1}, \alpha), \qquad j = 1, \ldots d-1 \tag{5.8.5}$$

define in $\Gamma_{d-1}(\frac{1}{2}\delta)$ *a family of surfaces* σ_α *parameterized by* $\alpha \in (\alpha_c - \varepsilon, \alpha_c + \varepsilon)$ *which are locally invariant, locally attractive, and tangent to the plane* $\eta_{s+1} = \ldots = \eta_{d-1} = 0$ *in a sense analogous to Eqs. (5.6.9)-(5.6.11). The tangency can be measured as in Eq. (5.6.11) in terms of an a priori given constant* $C > 0$.
(ii) Now assume that $s = 2$ *and that* $\lambda_1(\alpha) = \overline{\lambda_2(\alpha)}$ *is the eigenvalue of* L_α . *with largest modulus for all* $\alpha \in J$ *and that for* $\alpha = \alpha_c \in J$ *it is* $|\lambda_1(\alpha_c)| = 1$, $(\frac{d}{d\alpha}|\lambda_1(\alpha)|)_{\alpha=\alpha_c} > 0$, $\mathcal{I}\mathrm{m}\,\lambda_1(\alpha_c) \neq 0$ *and* $\lambda_1(\alpha_c)^h \neq 1$ *for* $h = 1, 2, 3, 4, 5$. *Suppose that the vague attractivity of* $\mathbf{0}$ *near* α_c *takes place because a condition analogous to Eq. (5.5.25),* $\overline{\gamma}_{\alpha_c} < 0$ *holds. Finally, assume that* k *is large enough and* $\alpha - \alpha_c$ *is small enough. Then there is a set on* σ*, which we denote* τ_α*, invariant with respect the to action of* $\widehat{\Phi}_{\sigma,\alpha}$ *and homeomorphic to a circle for* $\alpha > \alpha_c$*. Such a set is the intersection between* σ *and a torus which is invariant for the solutions of Eq. (5.8.1) and attracts, exponentially fast, all the motions starting close enough to it.*

Observations.
(1) Hence, in a similar way, as the vaguely attractive stationary points may bifurcate, in some circumstances, growing into periodic orbits, the periodic orbits may bifurcate growing into two-dimensional tori.
(2) The proof of the above proposition is parallel to that of Propositions 11, §5.6 and 12, §5.7, and will not be discussed in detail (see problems at the end of this section).
We only mention that the assumptions on the eigenvalues, at $\alpha = \alpha_c$, are needed to be able to put the transformation into a normal form analogous to Eqs. (5.7.4) and (5.7.15), thus allowing us to formulate a vague attractivity condition like Eq. (5.5.25).
(3) Proposition 14, together with Propositions 7-13 and the problems at the end of the §5.4-§5.8, provide a quite general theory of the stability of the vaguely attractive stationary points and periodic orbits and of their bifurcations, when the regularity class of the differential equation is high enough.
It then becomes natural to ask if it is possible to discuss in a similar fashion the theory of stability and bifurcations (following the loss of stability as a parameter a grows) of attractors or of more complex invariant sets.
"Unfortunately", such a question is very difficult, and it seems unsuited to be considered in too general a context. Only within classes of special cases, such a problem can be treated in some detail (e.g., in the case of the theory of the attractors "verifying the axiom A").[18] This is a theme of great interest, which seems to be connected with the theory of many phenomena more general than the ones of a purely mechanical nature, like the theory of turbulence which greatly stimulates research on this subject.
(4) As a comment on the generality of the theory of this and the preceding

[18] For a definition, see [45] and, also, [42] and [7] for detailed discussions of some problems (References).

sections, we must stress that the vague attractivity of a point or of an orbit near a critical value α_c is an interesting hypothesis, mainly for its elegant implications, but is far from being realized always (or even often). It often happens that simple systems of differential equations have stationary points or periodic orbits which are not vaguely attractive near a critical value α_c where they lose stability. In such cases, there is no general theory guiding the theoretical analysis of the attractors, and various phenomena are possible, like the "sudden" (i.e., for α just above α_c) transition to an asymptotic regime governed by attractors of a nature more complex than a stationary point or a periodic orbit or a two-dimensional torus. Such attractors may be located far from the attractor that lost stability.

In general attractors other than points, periodic orbits or tori run quasi-periodically are called "strange": this qualifies the impossibility of describing these attractors as simple objects, rather than qualifying a well-defined mathematical property.

To illustrate Observations (3), (4) and to get some feeling for how complicated the pattern of the bifurcations may be even for relatively simple differential equations (with quadratic nonlinearities "only"), we give a series of examples.

Some of the results quoted below may be obtained via the theory of the preceding section (like those relative to the stability of the stationary solutions, see §5.4 and §5.5 and the associated problems), possibly using a computer to estimate the eigenvalues of various stability matrices. However, most of the following results can only at present be obtained via the use of numerical experiments (usually fascinating). They should not be considered as mathematical statements but as empirical observations which may reveal themselves only as first rough approximations to the phenomena that the same nonlinear differential equations may show if studied more carefully.

We leave to the reader, as interesting practical work, the task of checking the following statements analytically (when possible) or numerically (if he has access to a computer: for the purpose the software in Appendix U can be used in a first approach).

5.8.1 A. Example 1: The "Lorenz Model".

Analytically, this is a system of equations that the reader can interpret as equations of motion of a gyroscope subject to suitable forces (following a scheme like the one in §5.1). The equations are

$$\dot{x} = -\sigma x + \sigma y, \quad \dot{y} = -\sigma x - y - xz, \quad \dot{z} = -bz + yx - \alpha \tag{5.8.6}$$

$\alpha = 10$, $b = \frac{8}{3}$. This system admits a "symmetry" group, i.e. a group of maps transforming solutions into solutions: namely the two elements group consisting in the maps $x' = \sigma x, y' = \sigma y, z' = z$, $\sigma = \pm 1$. The following items describe the structure of the attractors.

(1) For $0 \le \alpha < \alpha_c^1 = -b(\sigma - 1)$, there is just one stationary point that can be shown to be globally attractive for a small enough. It is locally stable and the eigenvalues of the Lyapunov matrix have a negative real part, $\forall \alpha \in (0, \alpha_c^1)$, and, numerically, it appears to

be globally attractive all the way up to α_c^1. The stationary point is stationary for all α_c but is unstable for $\alpha > \alpha_c^1$. It is

$$x = y = 0, \qquad z = -\frac{\alpha}{b} \tag{5.8.7}$$

and the symmetry maps leave it invariant.
(2) For $\alpha_c^1 < \alpha < \alpha_c^2 = 2\sigma\, b\, \frac{1+\sigma}{\sigma-1-b} = \frac{1760}{19} \simeq 92.63$, the preceding point undergoes a bifurcation, losing stability in one real direction but remaining vaguely attractive and it bifurcates in two locally stable stationary solutions which are mapped into each other by the symmetry maps. Such solutions exist for all $\alpha > \alpha_c^1$, but lose stability for $\alpha > \alpha_c^2$.
From a numerical point of view, a randomly chosen initial datum is attracted by one of the above two stationary solutions. The solutions are

$$x = y = \pm\sqrt{\alpha - b(\sigma+1)}, \qquad z = \sigma - 1. \tag{5.8.8}$$

One should not think, however, that the possible asymptotically different motions consist of the three points of Eqs. (5.8.7) and (5.8.8). For instance, for $\alpha < \alpha_c^2$ and close to it, there are some unstable periodic orbits, as can be rigorously shown.[19]
The reason why such asymptotic motions cannot be seen by sampling randomly the initial data space is that they form a set of zero Lebesgue measure.
(3) For $\alpha > \alpha_c^2$ the points of Eq. (5.8.8) lose stability. Such loss of stability takes place in two complex-conjugate directions because two complex-conjugate non real eigenvalues $(\pm\, i\, \sqrt{\alpha_c^2})$ of the stability matrix cross the imaginary axis from left to right.
However, although the fixed points in Eqs. (5.8.8) still exist for all $\alpha > \alpha_c^1$, they are not vaguely attractive for α near α_c^2. Hence, one cannot apply the Hopf bifurcation theorem to infer the existence of a bifurcation into periodic orbits of each of the points of Eq. (5.8.8). In fact, a strange attractor shows up here, see Fig. 5.8.

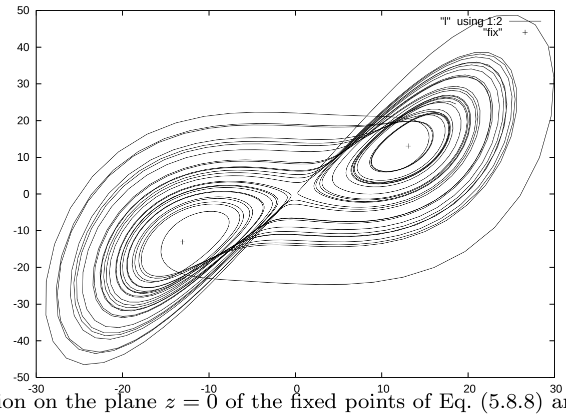


Figure 5.8 Projection on the plane $z = 0$ of the fixed points of Eq. (5.8.8) and of a motion corresponding to a given initial datum randomly chosen; $\alpha = 200$. The motion is not periodic. The marks are the projections of the (unstable) fixed points.

It exists up to $\alpha \simeq 230$, disappearing occasionally only for some small intervals of α when it is replaced by some stable periodic orbits: see Fig. 5.9, 5.10

[19] Applying Problem 16, §5.5, p.408, to either of the Eqs. (5.8.8) near α_c^2, one computes the vague-attractivity indicator of Proposition 12, Eq. (5.7.5) and shows that it has the "wrong sign", $\overline{\gamma}.0$, and then one applies Problem 9, §5.7.1, p. 440.

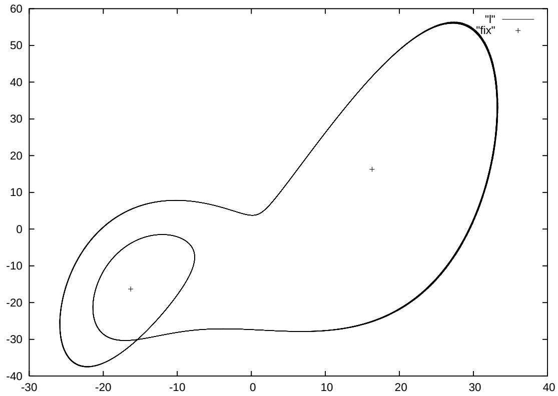


Figure 5.9 x, y projection of a periodic orbit relative to the case $\alpha = 340$. The other periodic orbits that can be experimentally found turn out to be related to the above by the transformation $x \to -x, y \to -y, z \to z$ which is a symmetry of the equation.

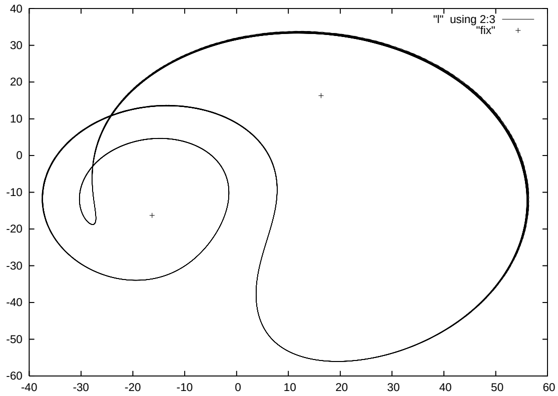


Figure 5.10 y, z projection of the orbit in Fig. 5.9.

(4) For large α the strange attractor disappears and is replaced by attractors consisting of periodic orbits, as it appears from numerical experiments. The existence of some stable periodic orbits can be proven rigorously for a large (see [41]).

5.8.2 B. Example 2: Navier-Stokes equations on a two-dimensional torus with a five mode truncation.

This is an example in which there are nice Hopf bifurcations. It is, however, more complicated than Example 1. It could also be interpreted mechanically as a system of two coupled rigid bodies with a rather strange looking coupling. However, this mechanical interpretation does not seem to be particularly useful, and we do not discuss it. The physical origin of the model has to be searched for in the theory of fluids. The equations are

$$
\begin{aligned}
\dot\gamma_1 &= -\,2\gamma_1 + 4\gamma_2\gamma_3 + 4\gamma_4\gamma_5,\\
\dot\gamma_2 &= -\,9\gamma_2 + 3\gamma_1\gamma_3,\\
\dot\gamma_3 &= -\,5\gamma_3 - 7\gamma_1\gamma_2 + \alpha,\\
\dot\gamma_4 &= -\,5\gamma_4 - \gamma_1\gamma_5,\\
\dot\gamma_5 &= -\,\gamma_5 - 3\gamma_1\gamma_4.
\end{aligned}
\tag{5.8.9}
$$

The equations are symmetric under a four elements symmetry group, namely $\gamma_1 \to \varepsilon\gamma_1$, $\gamma_2 \to \varepsilon\gamma_2, \gamma_3 \to \gamma_3$, $\gamma_4 \to \sigma\gamma_4$, $\gamma_5 \to \varepsilon\sigma\gamma_5$ with $\varepsilon, \sigma = \pm 1$.

(1) For α small, the obvious stationary solution, existing $\forall\, \alpha > 0$,

$$
\gamma_1 = \gamma_2 = \gamma_4 = \gamma_5 = 0, \quad \gamma_3 = \frac{\alpha}{5}
\tag{5.8.10}
$$

is stable and globally attractive [this could be proved along the lines of the proof of Eq. (5.2.12) in Proposition 4, §5.2, p.371]. By the Lyapunov criterion, it remains stable up to $\alpha_c^1 = 5\sqrt{\frac{3}{2}}$. Up to this value it appears, numerically, that it is a global attractor.

(2) Near α_c^1, Eq. (5.8.10) is vaguely attractive and loses stability in one real direction, generating two stable attractive solutions (5.8.11), mapped into each other by the symmetries

$$
\begin{aligned}
\gamma_1 =& \varepsilon\sqrt{\frac{\sqrt{6}}{7}}\,\sqrt{(\alpha-\alpha_c^1)}, \quad \gamma_3 = \sqrt{\frac{3}{2}}\\
\gamma_2 =& \varepsilon\sqrt{\frac{1}{7\sqrt{6}}}\,\sqrt{(\alpha-\alpha_c^1)}, \quad \gamma_4 = \gamma_5 = 0, \quad \varepsilon = \pm 1.
\end{aligned}
\tag{5.8.11}
$$

Such solutions exist for all $\alpha > a_c^1$ and, numerically, they seem to be globally attractive as long as they are locally stable: this means that randomly chosen initial data are attracted by either of them, see the comment to the point (2) of the Example 1 above.

They lose stability for $\alpha = \alpha_c^2$:

$$
\alpha_c^2 = \frac{80}{9}\sqrt{\frac{3}{2}}
\tag{5.8.12}
$$

The stability loss takes place in just one real direction again and, again, each of them bifurcates into two new stable solutions which are locally attractive for $\alpha \in (\alpha_c^2, \alpha_c^3)$, but persist for all $\alpha > \alpha_c^2$. If $\varepsilon, \sigma = \pm 1$

$$
\begin{aligned}
\gamma_1 =& \varepsilon\sqrt{\frac{5}{3}}, \quad \gamma_2 = \varepsilon\alpha\frac{3}{80}\sqrt{\frac{5}{3}}, \quad \frac{9}{80}\alpha,\\
\gamma_3 =& \frac{\sigma}{3}\sqrt{(\frac{9}{80}\alpha)^2 - \frac{3}{2}}, \quad \gamma_5 = -\sigma\sqrt{(\frac{9}{80}\alpha)^2 - \frac{3}{2}}\sqrt{\frac{5}{3}},
\end{aligned}
\tag{5.8.13}
$$

and $\alpha_c^3 = 22.8537\ldots$. The four points are mapped into each other by the symmetry group elements.

At $\alpha = \alpha_c^3$, Eqs. (5.8.13) lose stability in two complex directions and, apparently, they remain vaguely attractive. In fact, one can easily find, numerically, that in their vicinity there is a stable periodic orbit, as if a Hopf bifurcation had taken place (in principle, one could even check rigorously whether the vague-attractivity indicator $\overline{\gamma}$ is negative, as it probably is). The symmetry implies that the periodic orbits will be several: each bifurcating at $\alpha = \alpha_c^3$ from one of the four fixed points that become unstable. One of them is drawn in Fig. 5.11.

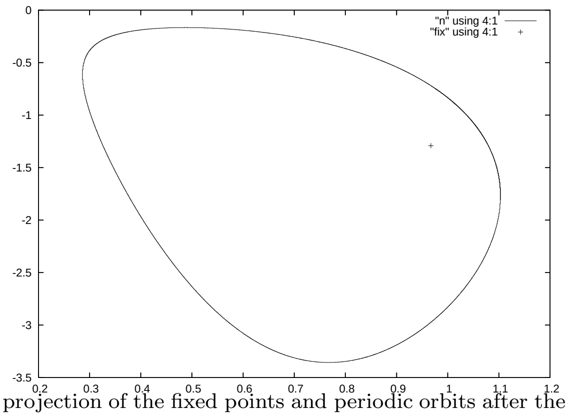


Figure 5.11 $\gamma_4 - \gamma_1$ projection of the fixed points and periodic orbits after the bifurcation in which the points of Eq. (5.8.13) lose stability ($\alpha = 28$). Eq. (5.8.9) has a fourfold symmetry ($\varepsilon, \varepsilon = \pm 1$) which can be used to generate three other orbits and fixed points symmetric to the one in the picture by applying the symmetry transformations mentioned in the text.

The structure of the motions for $\alpha > \alpha_c^3$ is quite fascinating. At various values $\alpha_c^{4,1}$, $\alpha_c^{4,2}$, $\alpha_c^{4,3}, \ldots$ there appear new periodic orbits bifurcating from the preceding ones because the latter lose stability in one real direction, with the stability matrix of the Poincare transformation showing the largest eigenvalue crossing the unit circle through -1.

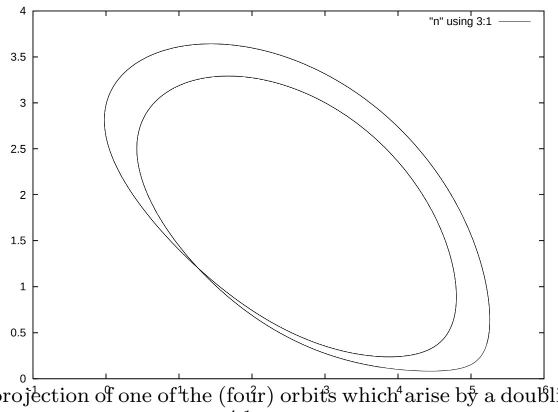


Figure 5.12 $\gamma_1 - \gamma_4$ projection of one of the (four) orbits which arise by a doubling bifurcation from one of the orbits of Fig. 5.11 for $\alpha = \alpha_c^{4.1}$ ($\alpha = 28.60$). the other three doubled periodic orbits are obtained from this one by the symmetry operations.

Such cases, although not contemplated in Proposition 14, can nevertheless be theoretically treated under suitable vague-attractivity assumptions, and their theory predicts that the periodic orbit "doubles", doubling also its period,[20] see also Problems 10-13 for §5.8.

[20] This can easily be understood intuitively by arguing as in the Observation (6) to Proposition 12, p.431. Write the Poincaré map as $\widehat{\Phi}_{\sigma,\alpha}(x) = (-1 - (\alpha - \alpha_c))x + p(x,\alpha)$, assuming that $x\, p(x,\alpha)/x^4 \xrightarrow[x\to 0]{} \overline{\gamma} < 0$. One easily finds that there are two points $x_{+,\alpha}, x_{-,\alpha} \simeq \pm\sqrt{-\overline{\gamma}^{-1}(\alpha - \alpha_c)}$ *mapped into each other* by $\widehat{\Phi}$. This means that the orbit "doubles".

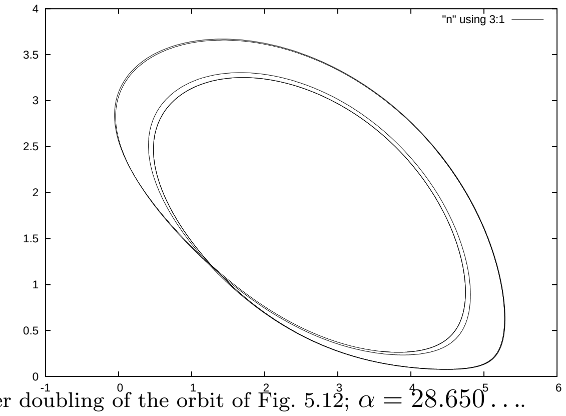


Figure 5.13 Further doubling of the orbit of Fig. 5.12; $\alpha = 28.650\ldots$.

The sequence of such bifurcations seems to be infinite and has been observed until the period has reached approximately 2^5 times the initial value. The accumulation point $\lim_{n\to\infty} \alpha_c^{4,n}$, as experimentally measured by a computer, seems to be $\alpha_c^{4,\infty} = 28.6681\ldots$.

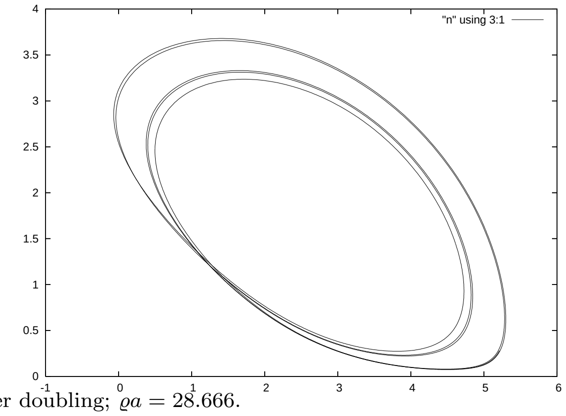


Figure 5.14 Further doubling; $\varrho a = 28.666$.

For $\alpha = \alpha_c^{50} = 28.663\ldots$, there appears a *new fourfold family* of periodic orbits (symmetric of each other under the symmetry group) that in the narrow interval $\alpha \in [\alpha_c^{5,0}, \alpha_c^{4,\infty}]$ coexists with the preceding ones, although they are also stable. A randomly chosen initial datum is attracted by one of the stable orbits of the two families, i.e. by one of the eight stable periodic orbits.

One of the new orbits (quite different in structure and location in phase space) is drawn in Fig. 5.15.

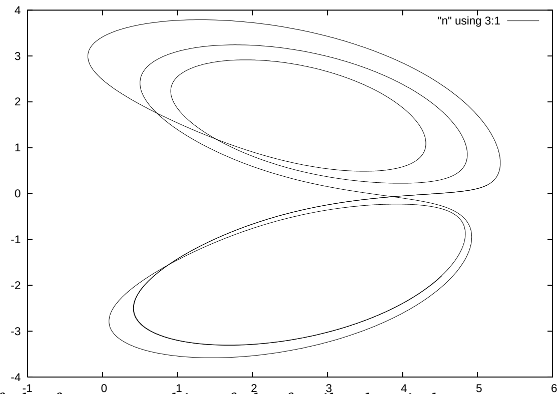


Figure 5.15 One of the four new orbits of the family that is born at $\alpha = \alpha^{5,0} = 28.663$ for $\alpha = 28.663$. The other four orbits are obtained from this by transforming it with the symmetries of the equation.

As a grows beyond $\alpha_c^{5,0}$ these new orbits also undergo the same fate, doubling after losing stability into a double orbit at $\alpha = \alpha_c^{5,1}$ which, in turn, doubles into a double orbit at $\alpha_c^{5,2}$, etc. "indefinitely" with an accumulation point at $\alpha = \alpha_c^{5,\infty} = 28.7201\ldots$. An example of the bifurcation is drawn in Fig. 5.16.

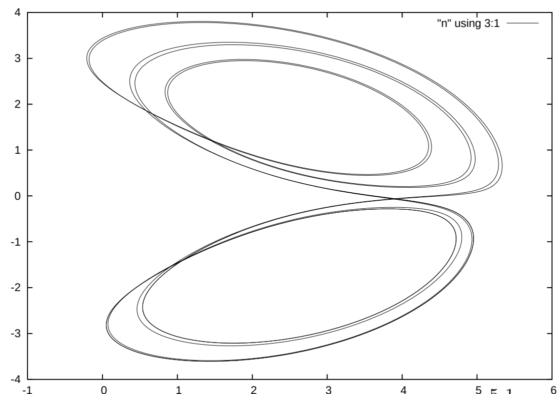


Figure 5.16 Figure 5.16. Doubling of the orbit in Fig. 5.15 for $\alpha > \alpha_c^{5,1}$; $\alpha = 28.710$.

For $\alpha > a_c^{5,\infty}$, it seems that the motion is asymptotically described by a strange attractor up to $\overline{\alpha}_c \simeq 34$ with the exception of at least one small interval of values of α, very small, where asymptotic behavior is again ruled by some periodic orbits which, as a grows, lose stability "again through -1" doubling in period infinitely many times. See the γ_3, γ_1 projection of the attractor for $\alpha = 31$.

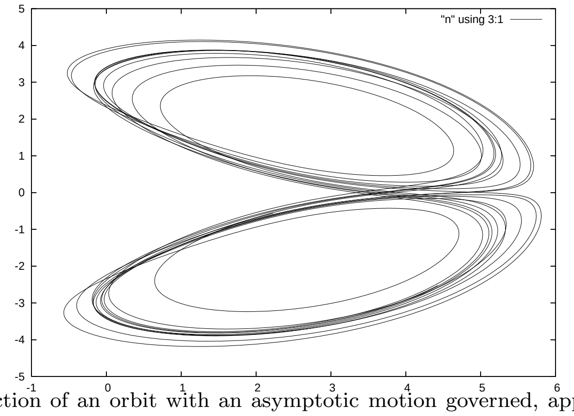


Figure 5.17 Projection of an orbit with an asymptotic motion governed, apparently, by a strange attractor; $\alpha = 31$.

After $\overline{\alpha}_c$, the motion seems to be governed by periodic and globally attractive orbits whose period and shape vary regularly with α (as before, here global "numerical" attractivity means that if the initial datum is randomly chosen, it converges to one of the above periodic motions). An example is drawn in Fig. 5.18

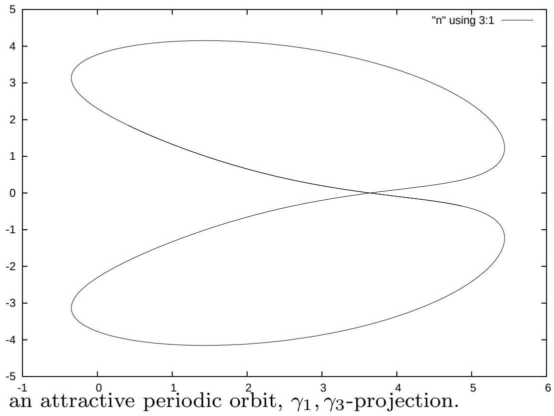


Figure 5.18 $\alpha = 34$; an attractive periodic orbit, γ_1, γ_3-projection.

We stress that the adjective "numerical", referred to some properties of the solutions, means that such properties come out of a computer-assisted study and that they are not mathematically rigorous.

Another exceptionally interesting and marvelous property of the above sequences of bifurcations is that, numerically, the sequences

$$\frac{\alpha_c^{4,n+1} - \alpha_c^{4,n}}{\alpha_c^{4,n} - \alpha_c^{4,n-1}}, \quad \frac{\alpha_c^{5,n+1} - \alpha_c^{5,n}}{\alpha_c^{5,n} - \alpha_c^{5,n-1}}$$

seem to converge to a limit ϱ^{-1} which is $\varrho^{-1} \simeq 4.67$. This is a numerical value which is conjectured, "Feigenbaum conjecture", to be "universal", i.e., independent of the particular differential equations giving rise to stable periodic orbits which successively grow out of doubling bifurcations when one of them, stable at a given value of α, loses stability as a grows, giving rise to a stable doubled orbit, [15].

However, it is an open problem to formalize in satisfactory generality and to give manageable sufficient conditions for a proof of the validity of this fascinating conjecture which seems to be verified in several cases studied numerically (and different from the above-considered ones). Recently, considerable progress in this direction has been achieved (see [9], and [8], and [30]).

The structure of the just discussed bifurcations is illustrated by Figs. 5.11-5.18, representing projections on several planes of trajectories of Eq. (5.8.11).

5.8.3 C. Example 3: Navier-Stokes equations on a two-dimensional torus with seven modes.

A system exhibiting periodic orbits bifurcating into two-dimensional tori along the scheme suggested by Proposition 14 is the following:

$$\begin{aligned}
\dot\gamma_1 &= -\,2\gamma_1 + 4\sqrt{5}\gamma_2\gamma_3 + 4\sqrt{5}\gamma_4\gamma_5,\\
\dot\gamma_2 &= -\,9\gamma_2 + 3\sqrt{5}\gamma_1\gamma_3,\\
\dot\gamma_3 &= -\,5\gamma_3 - 7\sqrt{5}\gamma_1\gamma_2 + 9\gamma_1\gamma_7 + \alpha,\\
\dot\gamma_4 &= -\,5\gamma_4 - \sqrt{5}\gamma_1\gamma_5,\\
\dot\gamma_5 &= -\,\gamma_5 - 3\sqrt{5}\gamma_1\gamma_4 - 5\gamma_1\gamma_6,\\
\dot\gamma_6 &= -\,\gamma_6 + 5\gamma_1\gamma_5,\\
\dot\gamma_7 &= -\,5\gamma_7 - 9\gamma_1\gamma_3,
\end{aligned} \tag{5.8.14}$$

which can be discussed in a similar way as that of Example 2.

The structure of the bifurcations and attractors is considerably more complicated and interesting. We do not discuss it in detail, feeling that Figs. 5.19-5.23 will, by themselves, excite the reader's curiosity and will stimulate him to read some original papers on the profound theory of Feigenbaum, [15], and on Example 3 as well as on Examples 1 and 2 (see [15], [18], [19], [17],[47]).

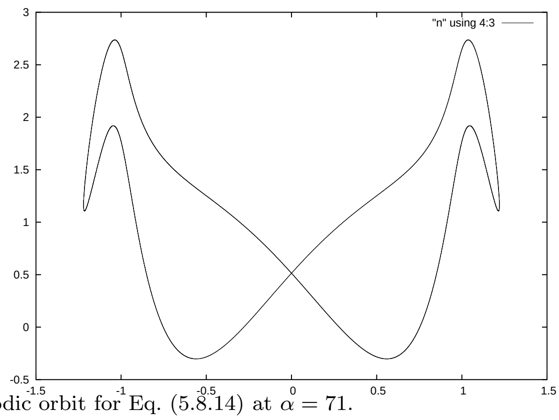


Figure 5.19 A periodic orbit for Eq. (5.8.14) at $\alpha = 71$.

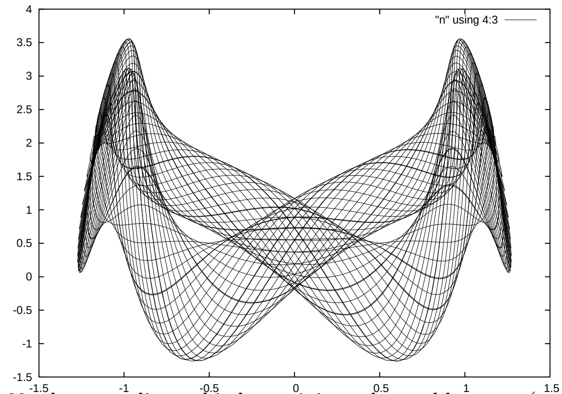


Figure 5.20 $\alpha = 71.60$; the preceding orbit has originated a stable torus (two dimensional) run quasi-periodically by the motions of Eq. (5.8.14), one of which is shown here.

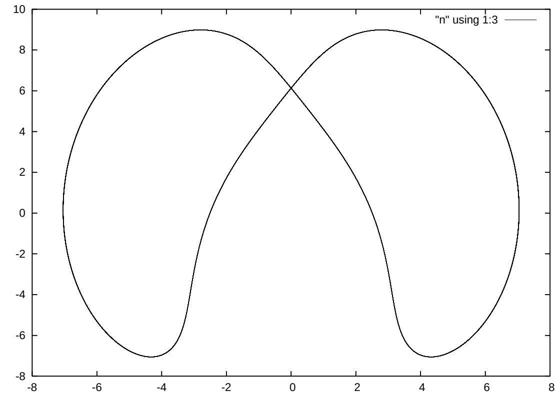


Figure 5.21 $\alpha = 190$; another stable periodic orbit.

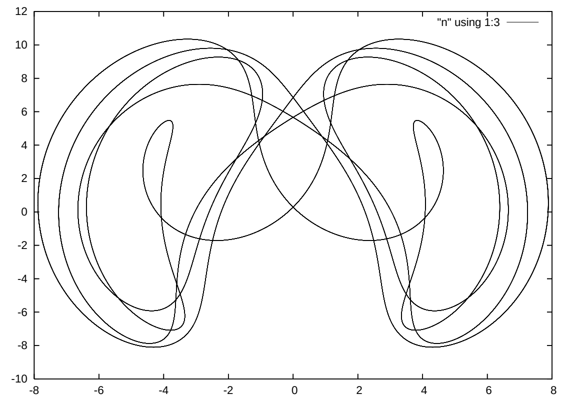


Figure 5.22$\alpha = 190$; another stable periodic orbit which coexists with that of Fig. 5.21. A randomly chosen initial datum, at this value of α is attracted either by the periodic motions

of Figs. 5.21 and 5.22 [or some of their images by the symmetries of Eq. (5.8.14)] or by the quasi-periodic motion which takes place on the torus of Fig. 5.23.

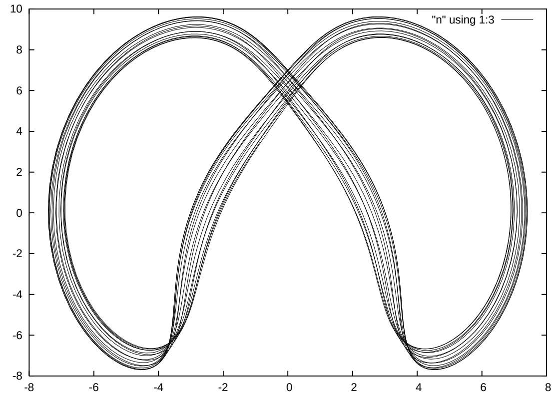


Figure 5.23 $\alpha = 195$: a stable two-dimensional torus run quasi-periodically by the motions of Eq. (5.8.14). This torus is an attractor apparently bifurcating from one of the periodic orbits in Fig. 5.21. Tori of dimension 2 can be quite easily identified by plotting a 2-dimensional section and checking that if can be fitted by a smooth closed curve: this can be done for instance for the torus in this figure.

All the equations of the above examples, as noted in Examples 1 and 2, can be interpreted as equations governing some strange systems of coupled rigid bodies, but they have been considered in the literature as equations approximating the differential equations describing the motion of simple fluids (like the "Euler" or the "Navier-Stokes" equations or the "thermo-fluidodynamics" equations). Their connection with the mechanics of rigid bodies is not surprising, however, if one notes that the classical fluid equations (Euler or Navier-Stokes equations) can be considered as equations describing infinitely many coupled rigid bodies (with very strange and, perhaps, mechanically unnatural coupling); this remark becomes clearer if one recalls that the equations of motion of the fluid bodies are usually derived by thinking of them as consisting of many small rigid bodies and applying to each of them the cardinal equations of mechanics.

We shall not further pursue the discussion of the models of dissipative systems and of their stability theory. This is a subject under current intense investigations, and the contents of §5.1-§5.8 provide some introduction to the literature.

5.8.4 Problems and Complements

1. Let $\Phi_0 \in C^\infty(\mathcal{R}^d)$ be a map of $\mathcal{R}^d$ into itself with the origin as a fixed point. Write $\mathbf{x}' = \Phi(\mathbf{x})$ as

$$\mathbf{x}' = L\mathbf{x} + \mathbf{F}(\mathbf{x}).$$

where L is a $d \times d$ matrix and $\mathbf{F}$ has a second-order zero at the origin. Suppose that the eigenvalues of L are pairwise distinct. Show that there is a linear change of coordinates that allows us to put the above map into the form

$$\begin{aligned} x_1^{(j)'} &= (\mathcal{R}\mathrm{e}\,\lambda_j)x_1^{(j)} - (\mathcal{I}\mathrm{m}\,\lambda_j)x_2^{(j)} + F_1^{(j)}(\mathbf{x}) \\ x_2^{(j)'} &= (\mathcal{I}\mathrm{m}\,\lambda_j)x_1^{(j)} - (\mathcal{R}\mathrm{e}\,_j)x_2^{(j)} + F_1^{(j)}(\mathbf{x}) \\ x_1^{(j)'} &= \lambda_h + F^{(h)}(\mathbf{x}) \end{aligned}$$

with $j = 1, \ldots, s$, $h = 2s+1, \ldots, d$, where $\lambda_1, \ldots, \lambda_s$ are the s complex non real eigenvalues of L and $\lambda_{2s+1}, \ldots, \lambda_d$ are the $(d-2s)$ real eigenvalues of L; $F_1^{(j)}, F_2^{(j)}, \ldots F^{(h)}$ have a second-order zero at the origin $\mathbf{0}$. (*Hint:* Proceed as in the proof of Proposition 7, p.393, §5.5.)

2. In the context of Problem 1, suppose that $d = 2, \lambda = \lambda_1 = \{\text{complex non real}\}$. Let $z = x_1^{(1)} + ix_2^{(1)}$. Show that the map can be written as a map of $\mathcal{C}$ into itself:

$$z' = \lambda z + F(z, \overline{z}),$$

where F has a second-order zero at $z = 0$.

3. Show that if $\lambda^2 \neq 1, \lambda \neq 0$, the map in Problem 2 can be written in a new coordinate system as

$$\zeta' = \lambda\zeta + N(\zeta, \overline{\zeta}),$$

where N has a third-order zero at the origin $\zeta = 0$. (*Hint:* Proceed as in the proof of Proposition 8, p.396, §5.5, i.e., write $F(z, \overline{z}) = a_2 z^2 + a_1 z\overline{z} + a_0\overline{z}^2 + \widetilde{N}(z, \overline{z})$ with $\widetilde{N}$ having a third-order zero at $z = 0$. Change variables near $z = 0$ as $\zeta = z + A_2 z^2 + A_1 z\overline{z} + A_0\overline{z}^2$ and choose the A's in order to eliminate the second-order terms from the map in the new coordinates.)

4. Show that if $\lambda^4 \neq 1, \lambda \neq 0$ the map in Problem 3, of $\mathcal{C}$ into itself,

$$\zeta' = \lambda\zeta + N(\zeta, \overline{\zeta}),$$

with N having a third-order zero at$\zeta = 0$, can be put into the form

$$z' = \lambda z + b\, z|z|^2 + Q(z, \overline{z})$$

with Q having a fourth-order zero at $z = 0$, using a change of variables (near the origin) of the form: $z = \zeta + A_3\zeta^3 + A_2\zeta^2\overline{\zeta} + A_1\zeta\overline{\zeta}^2 + A_0\overline{\zeta}^3$.

5. In the context of Problem 4, show that the map can also be written as

$$z' = \lambda z\, e^{b|z|^2 + \widetilde{Q}(\varrho, \theta)}$$

near $z = 0$, where $z = \varrho\, e^{i\,\theta}$ and Q is a C^∞ function of $(\varrho, \theta) \in \overline{\varrho} \times \mathcal{T}^1$ with a third-order zero at the origin of the ϱ variable.

6. Consider the map defined as follows: let $z = \varrho e^{i\theta}$ and

$$z' = \lambda(\alpha)\, z\, e^{b(\alpha)|z|^2 + \widetilde{Q}(\varrho, \theta, \alpha)} \equiv \varPhi_\alpha(z)$$

where $\widetilde{Q} \in C^\infty([0, p) \times \mathcal{T}^1 \times (-a, a))$, $\lambda, b \in C^\infty((-a, a))$, $|\lambda(0)| = 1$, and $\widetilde{Q}$ has a third-order zero at $\varrho = 0$, for all $\theta \in \mathcal{T}^1$, for all $\alpha \in (-a, a)$. Show that the origin is vaguely attractive near zero if $\mathcal{R}\mathrm{e}\, b(0) < 0$. (*Hint:* If $\mathcal{R}\mathrm{e}\, b(0) < 0$ the origin is attractive for $\alpha = 0 \ldots$.)

7.* Let $\lambda(\alpha) = e^{\alpha+ib(\alpha)}$ and, in the context of Problem 6, let $\mathcal{R}\mathrm{e}\, b(0) < 0$. Show that the maps Φ_α have an attractive invariant set of approximate equation $|z| = \sqrt{\frac{\alpha}{-\mathcal{R}\mathrm{e}\, b(\alpha)}}$ for $\alpha > 0$ small. (*Hint:* Proceed as in the analysis of the Hopf theorem, performing the analogous steps and estimates.) Actually (but this is more difficult than the above problem), the invariant set is a curve homeomorphic to a circle. The proof of this could be achieved by writing the equation of the unknown curve as

$$z(\theta) = \sqrt{\frac{-\alpha}{\mathcal{R}\mathrm{e}\, b(\alpha)}}\,(1+\varepsilon(\theta))e^{i\theta}$$

and trying to determine $\varepsilon(\theta)$ by writing the condition that the above curve is Φ_α invariant, i.e.,

$$\begin{aligned}\theta' =& \theta + \beta(\alpha) - \alpha\frac{\mathcal{I}\mathrm{m}\, b(\alpha)}{\mathcal{R}\mathrm{e}\, b(\alpha)}(1+\varepsilon(\theta))^2 + (\sqrt{\alpha})^3\overline{Q}(1+\varepsilon(\theta),\theta,\alpha),\\ 1+\varepsilon(\theta') =& (1+\varepsilon(\theta))e^{-\alpha[(1+\varepsilon(\theta))^1-1]} + (\sqrt{\alpha})^3\overline{Q}_1(1+\varepsilon(\theta),\theta,\alpha),\end{aligned}$$

where Q, Q_1 are smooth functions of their three arguments. The equation can be solved recursively. The proof, however, is not really straightforward (see [29]).

8. Prove the first part of Proposition 14 for $d = 2, s = 1$. (*Hint:* Proceed as in the proof of Proposition 11, §5.6, p.411. Here the transformation in Problem 6 plays the role played there by the equation in normal form.)

9. Prove the second part of Proposition 14 for $d = 2$, assuming that the invariant set of Problem 7 is actually homeomorphic to a circle and making use of Problems 2-7 for the reduction to normal form.

10. Consider the C^∞ map Φ of $\mathcal{R}^1$ into itself:

$$x' = \Phi(x) = \lambda x + g(x)$$

with $g \in C^\infty(\mathcal{R})$ having a second-order zero at the origin. Show that if $\lambda \neq 0, 1$, there is a change of variables transforming the above map into a new one having the form

$$\xi' = \lambda\xi + \xi^3\gamma(\xi)$$

with γ in $C^\infty(\mathcal{R})$, for ξ near 0. (*Hint:* Let $g(x) = \overline{g}x^2 + \widetilde{g}(x)$ with $\widetilde{g}$ having a third-order zero at the origin. Set $\xi = x + Gx^2$ and find a suitable G.)

11. In the context of Problem 10, show that if

$$x' = -(1+\alpha)x + x^3\gamma(x,\alpha) \stackrel{def}{=} \Phi_\alpha(x)$$

is a family of maps of class C^∞ parameterized by α with $\gamma \in C^\infty(\mathcal{R}^2)$, $\gamma(0,0) > 0$ then there exist two points $x_+(\alpha), x_-(\alpha)$, for $\alpha > 0$ small, such that

$$\Phi_\alpha(x_+(\alpha)) = x_-(\alpha), \qquad \Phi_\alpha(x_-(\alpha)) = x_+(\alpha),$$

i.e., constituting a period 2 orbit ("doubling bifurcation"). Furthermore, show that by the Lyapunov criterion, such an orbit is stable and attractive. (*Hint:* Use the implicit function theorem to find $x_+(\alpha)$, say, as a root of $\Phi_\alpha^2(x) = x$. Prove the stability by applying the criterion of Lyapunov, Proposition 13, §5.8, p.441, to $x_+(\alpha)$ and to the map Φ_α^2.)

12. Consider a map $\mathbf{x}' = \boldsymbol{\Phi}(\mathbf{x},\alpha)$ of $\mathcal{R}^d$ into itself, parameterized by $\alpha \in \mathcal{R}$. Let $\boldsymbol{\Phi} \in C^\infty(\mathcal{R}^d \times R)$, let the origin be a fixed point of the map, for all α near zero, and let $L(\alpha)$ be its stability matrix. Suppose that for $\alpha \in (-a,a)$, all the eigenvalues of $L(\alpha)$ are pairwise

distinct and such that $|\lambda_1(0)| = 1 > \nu > |\lambda_2(0)|, \ldots, |\lambda_d(0)|$, with $\nu < 1$.
Using the attractive manifold theorem described in the first part of Proposition 14, p.442, and Problem 11, show that if $\lambda(\alpha) = -1 - \alpha$ then the origin undergoes a "period doubling bifurcation" as α grows through zero (in the sense of Problem 11). (*Hint:* Use the attractive manifold theorem to reduce the problem to a one dimensional problem and then apply Problems 10 and 11.)

13. Prove that Problem 12 implies that if $\boldsymbol{\Phi}(\mathbf{x}, \alpha)$ is (an arbitrary extension of) the Poincaré map for a periodic orbit of a one-parameter family of differential equations in $\mathcal{R}^{d+1}$, then the periodic orbit bifurcates to a stable (exponentially attractive) periodic orbit, as α grows through 0, with roughly a double period.

14. Study the map $x' = 4\alpha\, x(1-x)$, $x \in \mathcal{R}$, and show that $[0,1]$ is an invariant set if $\alpha \in [0,1]$. Find the first bifurcation of the fixed points $x = 0$ and $x = x_\alpha > 0$, $x_\alpha = 1 - \frac{1}{4\alpha}$ (consider the latter only for $\alpha > \frac{1}{4}$). Show that in some sense x_α grows out of a bifurcation of $x = 0$; while when x_α loses stability, it undergoes a doubling bifurcation in the sense of Problem 11.

15. Consider the map Φ in Problem 14 for $\alpha = 1$, restricted to $[0,1]$. Show that the change of variables $y = \frac{2}{\pi} \arcsin \sqrt{x}$ transforms this map into the mapΨ:

$$\Psi : y \to \begin{cases} 2y & \text{if } 0 < y < \frac{1}{2}, \\ 2(1-y) & \text{if } \frac{1}{2} < y < 1. \end{cases}$$

Draw (roughly) the graph of Ψ^n and show that Ψ^n has (by inspection of the graph) 2^n fixed points which correspond to 2^n periodic points for Ψ. Deduce that Φ also has 2^n periodic points of period n (here the period is not necessarily minimal).

16. Using Problem 15, show that Ψ and Φ have a dense set of periodic points. (*Hint:* Look at the graph of Ψ^n.)

17. Study the stability of the fixed points of the map of $\mathcal{R}^2 \to \mathcal{R}^2$ parameterized by α, b ("Henon's map"):

$$H(x,y) = (y - \alpha x^2 + 1, bx)$$

with b real and find whether one of its fixed points undergoes, for some fixed value of b, a doubling bifurcation as α grows using Problems 11 and 12.

18. Let $\mathbf{x} \to \boldsymbol{\Phi}(\mathbf{x})$ be a C^∞ map of the plane into itself which is invertible and area preserving (i.e., $\text{area}\, E = \text{area}\, \Phi^{-1}(E)$ for all measurable sets E). Which relation between the eigenvalues of the Lyapunov stability matrix of a fixed point follows as a consequence of the conservation of the area?

19. Same as Problem 18 for a volume-preserving map of $\mathcal{R}^d$ into itself.

20. In the context of Proposition 13, p.441, show that the eigenvalues of the stability matrix of a periodic orbit depend neither on the particular system of coordinates introduced on σ nor on the point $\boldsymbol{\xi}_0$ chosen on the orbit. They are "characteristic numbers" of the orbit itself. (*Hint:* This is a problem analogous to Problem 15, p.388, §5.4. The first statement is proven in exactly the same way. To prove the second, use the trajectories to "transfer" a system of coordinates on σ (through $\boldsymbol{\xi}_0$) into a system of coordinates on σ' (through $\boldsymbol{\xi}_0'$, etc).)

5.9 Stability in Conservative Systems: Introduction

...desinas ineptire
et quod perisse vides perditum ducas

Stability of Hamiltonian motions is a natural problem arising, perhaps for the first time, in the theory of the solar system, where it is still unsolved.

In nature there are many interesting systems which are "quasi-integrable" in the sense that their equations of motion differ, up to "quasi negligible" terms, from equations of motion of an integrable system.

A nice example is provided by the solar system which we consider via a model in which the solar mass M is $+\infty$, i.e., the Sun is a fixed point mass attracting the planets with a central force with potential energy inversely proportional to a planet distance and directly proportional to its mass. In the approximation in which the reciprocal attraction among the planets is neglected, it is clear that the solar system is described by as many Hamiltonian integrable systems as the number of planets (i.e., nine), one for each planet. In Chapter 4, §4.9.1 and §4.10.1, we saw that such Hamiltonian systems are integrable in the sense of Definitions 10 and 11, §4.8.1.

It is then attractive to think that the actual motion of the solar system is "close" to this idealized motion followed by nine independent planets.

Keeping for simplicity, the approximation that the Sun is a point mass fixed with respect to the fixed stars, we must compare the solutions of the following two systems of equations: $i = 1, \ldots, 9$,

$$m_i \ddot{\mathbf{x}}^{(i)} = -\frac{K m_i}{|\mathbf{x}^{(i)}|^2} \frac{\mathbf{x}^{(i)}}{|\mathbf{x}^{(i)}|} \tag{5.9.1}$$

$$m_i \ddot{\mathbf{x}}^{(i)} = -\frac{K m_i}{|\mathbf{x}^{(i)}|^2} \frac{\mathbf{x}^{(i)}}{|\mathbf{x}^{(i)}|} - \varepsilon \sum_{i \neq j} \frac{m_i m_j}{(\mathbf{x}^{(i)} - \mathbf{x}^{(j)})^2} \frac{(\mathbf{x}^{(i)} - \mathbf{x}^{(j)})}{|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}|}, \tag{5.9.2}$$

at least for initial data which, put into Eq. (5.9.1), give rise to trajectories on which $|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}|$, $i \neq j$, remains so large as to make the second term in the right-hand side of Eq. (5.9.2) small compared to the first. The constant ε is the universal gravitation constant, $m_1, \ldots, m_9$ are the masses of the nine main planets, $K = \varepsilon M_S$, where M_S is the Sun real mass; satellites, comets, asteroids, rings, etc. have been disregarded.

Choosing as time origin the flying away instant, the situation in which the solar system is initially found is, as is well known, such that the term in ε in Eq. (5.9.2) has a modulus quite a bit smaller than the term representing the Sun attraction. The first question, preliminary to the comparison between the solutions of Eqs. (5.9.1) and (5.9.2), is whether this situation remains unchanged as time goes by.

This property can be easily verified through the explicit solution of the various Kepler problems in the case of Eq. (5.9.1). Hence, this question is intimately related to the comparison between Eqs. (5.9.1) and (5.9.2).

From the general results of the theory of ordinary differential equations, it is evident that "close equations yield close solutions"; however, this closeness is *not uniform over time*. It does not, indeed, follow from the regularity theorems and the initial data and parameters dependence that close equations with close initial data produce solutions which stay close forever or solutions whose trajectories, as sets, remain close. The first possibility is almost always false.

One then asks if the corrections to the equations of motion (5.9.1) due to the presence of the term in ε in Eq. (5.9.2), though small, may lead to changes in the motions which, in the long run, result in a motion very different from the one foreseen in Eq. (5.9.1).

A priori, one could even consider "unthinkable" or undesirable catastrophic events, like interplanetary collisions or capture of a planet by the burning Sun.

Of course, one wishes to have analytic instruments for the solutions of Eq. (5.9.2) and of comparison with those of Eq. (5.9.1). The analysis should allow not only the exclusion of such catastrophes, but even to show that it is true, or essentially true, that the planets movements are described by Eq. (5.9.1). And furthermore that, if needed, one can compute or estimate the deviations between the motions of Eq. (5.9.1) and those of Eq. (5.9.2) with equal initial data at least for long times, i.e., of astronomical magnitude, long compared with the revolution periods of the various planets.

In other words, one wishes to use Eq. (5.9.1) for "rough" astronomical predictions and to have algorithms to compute the corrections at least for times of the order of magnitude of several thousand years.

That this is a delicate problem can be deduced from the fact that rough estimates, too pessimistic, of the errors lead to the conclusion that the reciprocal influence between the planets may become important within a few years.

For instance, the time necessary for a collision between two heavenly bodies of the size of Venus and Earth, assuming that at time zero they are standing still (relative to the fixed stars) at a distance $d(T,V)$, equal to the actual Earth-Venus maximal observed distance, could be estimated not longer than T_{coll} such that (accelerated motion estimate)

$$\frac{\varepsilon}{2}\frac{m_T+m_V}{d(T,V)^2}T_{coll}^2 = d(T,V) \;\Rightarrow\; T_{coll} \simeq 370\,\text{years}. \tag{5.9.3}$$

Hence, we see that even to establish some accurate predictions for times of a few centuries, a remarkable precision is needed, i.e., it is necessary to take into account the fact that the planets motion at the initial time is very far from a situation bound to a collision and that, obviously, the corrections to the motion described by Eq. (5.9.1), originated by the additional terms in ε in Eq. (5.9.2), are not always favorable to collisions (or escapes, etc.). Think of the two-body problem where the systematic attraction results only in providing a curvature to the trajectory. On the average, the effects favorable to catastrophic events may be much smaller or even totally absent with respect to the above pessimistic calculation.

This and similar problems, which may obviously be formulated for systems very different from the solar system (like harmonic oscillators with conservative anharmonic additional perturbing forces or, more generally, for systems "close" to integrable systems), are typical stability problems for conservative systems.

To the above problems, one adds analogous problems of stability of integrable systems perturbed by the addition, among the active forces, of external forces varying with simple time laws ("non autonomous Hamiltonian systems").

All of the above problems are much more difficult than one might imagine, perhaps naively. Only recently have some techniques apt to provide some answers been developed (and are being developed), although we are still quite far from a "satisfactory" theory even for very small perturbations.

The main result on this theme is the following theorem ("Kolmogorov-Arnold-Moser theorem") which we shall analyze in some particularly interesting cases in the §5.12. The reader who wishes to obtain deeper insights can consult [33, 34].

15 Proposition. *Consider a mechanical system in $\mathcal{R}^d$ with ℓ degrees of freedom, subject to conservative forces with potential energy $\Phi_0 \in C^\infty(\mathcal{R}^d)$ bounded from below and subject to ideal constraints.*
Suppose that the system is canonically integrable on some open set W of the phase space (see Definition 11, p.289, §4.8) and call H_0 its Hamiltonian.
If $I : W \longleftrightarrow V \times \mathcal{T}^\ell$ is the integrating transformation and if we set $(\mathbf{A}, \boldsymbol{\varphi}) = I(\mathbf{p}, \mathbf{q})$, the motion in $(\mathbf{A}, \boldsymbol{\varphi})$ coordinates is, by definition,

$$\widehat{S}_t(\mathbf{A}, \boldsymbol{\varphi}) = I(S_t(I^{-1}(\mathbf{A}, \boldsymbol{\varphi}))) = (\mathbf{A}, \boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{A})t), \tag{5.9.4}$$

where $\boldsymbol{\omega} = (\omega_1(\mathbf{A}), \ldots, \omega_\ell(\mathbf{A}))$ are ℓ pulsations corresponding to the ℓ prime integrals $\mathbf{A} = (A_1, \ldots, \mathbf{A}_\ell)$, and $\boldsymbol{\omega}(\mathbf{A}) = \partial_{\mathbf{A}} h_0(\mathbf{A})$ if $h_0(\mathbf{A}) = H_0(I^{-1}(\mathbf{A}, \boldsymbol{\varphi}))$ [$\boldsymbol{\varphi}$ independent because of the integrating character of I, see Observation (1), p.289]. Assume V to be bounded, and that the matrix

$$J_{ij} = \frac{\partial \omega_i}{\partial A_j} \tag{5.9.5}$$

has non vanishing determinant on all of V ("non isochrony" of the system).
Then, if $\Psi \in C^\infty(\mathcal{R}^d)$ is a uniformly bounded potential energy, the mechanical system with the same constraints but with an active force with potential energy

$$\Phi_0 + \varepsilon \Psi \tag{5.9.6}$$

has various remarkable properties which will be described calling $S_t^{(\varepsilon)}$ and $\widehat{S}_t^{(\varepsilon)}$, $t \in \mathcal{R}$, the transformations generating the motions, corresponding to Eq. (5.9.6) and to the given constraints, in the coordinates $(\mathbf{p}, \mathbf{q})$ and $(\mathbf{A}, \boldsymbol{\varphi})$, respectively. If $(\mathbf{A}, \boldsymbol{\varphi}) = I(\mathbf{p}, \mathbf{q})$, then

$$\widehat{S}_t^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi}) = I(S^{(\varepsilon)}(\mathbf{p},\mathbf{q})) \tag{5.9.7}$$

(and $\widehat{S}_t^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi})$ is only defined for those pairs $(\mathbf{A},\boldsymbol{\varphi})$ for which Eq. (5.9.7) makes sense).

(i) There is a subset $W^{(\varepsilon)} \subset W$ invariant for the transformations $S_t^{(\varepsilon)}$ and a map $\mathbf{F}^{(\varepsilon)|:W^{(\varepsilon)} \longleftrightarrow V^{(\varepsilon)}} \times \mathcal{T}^\ell$, $V^{(\varepsilon)} \subset V$, invertible and continuous, denoted

$$\mathbf{F}^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi}) = (\mathbf{a}(\mathbf{A},\boldsymbol{\varphi},\varepsilon),\boldsymbol{\Psi}(\mathbf{A},\boldsymbol{\varphi},\varepsilon)). \tag{5.9.8}$$

Furthermore, there is a continuous function $\boldsymbol{\Omega}^{(\varepsilon)} : W^{(\varepsilon)} \to \mathcal{R}^\ell$ such that

$$\mathbf{F}^{(\varepsilon)}(\widehat{S}_t(\mathbf{A},\boldsymbol{\varphi})) = (\,\mathbf{a}(\mathbf{A},\boldsymbol{\varphi},\varepsilon),\,\boldsymbol{\Psi}(\mathbf{A},\boldsymbol{\varphi},\varepsilon) + \boldsymbol{\Omega}^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi})\,t\,). \tag{5.9.9}$$

Therefore, the motions with initial datum in $W^{(\varepsilon)}$ can be thought of as rotations of an ℓ-dimensional torus.

(ii) The set $W^{(\varepsilon)} \subset W$ is generally only measurable in the sense of Lebesgue and not necessarily in the sense of Riemann, and its measure is such that

$$\frac{\text{volume}\, W^{(\varepsilon)}}{\text{volume}\, W} \xrightarrow[\varepsilon\to 0]{} 1. \tag{5.9.10}$$

(iii) The functions $(\varepsilon,\mathbf{A},\boldsymbol{\varphi}) \to \mathbf{F}^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi})$ can be extended to $C^{(k)}$ functions with arbitrary preassigned k on $(-1,1)\times V\times\mathcal{T}^1$ and the same can be said of the functions $(\mathbf{A},\boldsymbol{\varphi}) \to \boldsymbol{\Omega}^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi})$. Furthermore, such extensions have the property

$$\mathbf{F}^{(0)}(\mathbf{A},\boldsymbol{\varphi}) \equiv (\mathbf{A},\boldsymbol{\varphi}), \quad \boldsymbol{\Omega}^{(\varepsilon)}(\mathbf{A},\boldsymbol{\varphi}) \equiv \frac{\partial h_0(\mathbf{A})}{\partial \mathbf{A}}, \tag{5.9.11}$$

*(iv) If the original system is an analytic analytically integrable system and * is also analytic, then one can take $k = +\infty$ in (iii).*

Observations.

(1) This theorem tells us the sense in which perturbing an integrable system with proper pulsations "really" variable, see Eq. (5.9.5), i.e., "non isochronous", one obtains a system that can still be thought of as a system moving essentially in the same way as the unperturbed one, see Eq. (5.9.11).

(2) $W^{(\varepsilon)}$ can be thought of as foliated into invariant ℓ-dimensional tori with equations

$$(\mathbf{A},\boldsymbol{\varphi}) = (\mathbf{F}^{(\varepsilon)})^{-1}(\mathbf{a},\boldsymbol{\psi}), \qquad \boldsymbol{\psi} \in \mathcal{T}^\ell \tag{5.9.12}$$

parameterized by ℓ parameters $\mathbf{a} \in V^{(\varepsilon)}$. By Eq. (5.9.11), each of such tori is a slight deformation of the torus described by $\{\mathbf{a}\}\times\mathcal{T}^\ell$ in the original variables.

(3) Observation (2) is interpreted as saying that the foliation of the phase space into invariant tori (characteristic of the integrable systems) is, at least in the

canonically integrable anisochronous cases, preserved under small perturbations, provided one disregards a subset of phase space with small measure.
(4) The fact that $V^{(\varepsilon)}$ can only be shown to be Lebesgue measurable (and it probably cannot be chosen Riemann measurable) is quite unpleasant because it means that $W^{(\varepsilon)}$, although containing many points (for ε small) cannot be approximated by "nice sets" and, therefore, it becomes difficult to decide constructively whether a given point is or is not in $W^{(\varepsilon)}$. However, a little thought shows that (iii) partially solves this problem from a practical point of view.
(5) Note that a ℓ-dimensional torus in a 2ℓ-dimensional space[21] does not split the space $\mathcal{R}^{2\ell}$ into "interior" and "exterior" parts, unless $\ell = 1$. This is perhaps what makes clearer the incompleteness of the result (iii). In fact, a point beginning its motion in $W/W^{(\varepsilon)}$, i.e., outside the invariant tori, may "sneak" through the tori of the foliations very far from the vicinity of the unperturbed torus on which it would move if $\varepsilon = 0$. This phenomenon, called "Arnold diffusion", is not well understood, [36].

It would be nice to understand criteria sufficient for the existence of a Riemann-measurable set of initial data (possibly with positive measure) which does not undergo the Arnold diffusion. I.e., implying that, although only a Lebesgue measurable set of points in phase space moves essentially as if the perturbation were not present (i.e., quasi-periodically, on tori close to the unperturbed ones), there is a Riemann measurable set of points moving (perhaps not quasi periodically) close to the unperturbed tori located near the initial data.

In fact, this is what the numerical experiments sometimes seem to suggest. It can be rigorously proved for some non autonomous 1-degree of freedom systems (once the above theorem is extended, as can be done, to the non autonomous system with external periodic forces of Hamiltonian type) or for 2-degrees-of-freedom autonomous systems.

In such cases, however, the entire problem disappears as the motion takes place on a three-dimensional set (because, in the first case, the "phase space" is three dimensional (p, q, t) and in the second, although the phase space is four dimensional, the motion takes place on the three-dimensional surface of constant energy), and in $\mathcal{R}^3$ a two-dimensional torus has an interior and an exterior.
(6) The above theorem cannot be applied to perturbations of harmonic oscillators since the non isochrony condition of Eq. (5.9.5) is manifestly violated.

Nevertheless, if the ℓ pulsations $\boldsymbol{\omega}_0 = (\omega_1, \ldots, \omega_\ell)$ of the harmonic oscillator verify a "non resonance" or "Diophantine" condition: $\exists\, C < \infty, \alpha < \infty$ and

$$|\boldsymbol{\omega} \cdot \boldsymbol{\nu}|^{-1} \le C|\boldsymbol{\nu}|^{\alpha}, \qquad \forall\, \boldsymbol{\nu} \neq \mathbf{0} \tag{5.9.13}$$

[21] or $\mathcal{R}^{2\ell-1}$, if energy is taken into account, unless $\ell \le 2$.

where $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_\ell) \in \mathcal{Z}^\ell$ is an "integer vector", then Eq. (5.9.5) can be replaced by a condition on Ψ. Namely, if $f(\mathbf{A}, \boldsymbol{\varphi})$ is the function Ψ in the $(\mathbf{A}, \boldsymbol{\varphi})$ variables, $f(\mathbf{A}, \boldsymbol{\varphi}) = \Psi(I^{-1}(\mathbf{A}, \boldsymbol{\varphi})), (\mathbf{A}, \boldsymbol{\varphi}) \in V \times \mathcal{T}^\ell$, and if we define

$$f_{\mathbf{0}}(\mathbf{A}) = \frac{1}{(2\pi)^\ell} \int -\mathcal{T}^\ell f(\mathbf{A}, \boldsymbol{\varphi})\, d\boldsymbol{\varphi}, \tag{5.9.14}$$

and if the matrix $\frac{\partial^2(\mathbf{A})}{\partial A_i \partial A_j}$, $i, j = 1, \ldots, \ell$, has non vanishing determinant $\forall\, A \in V$, the theorem's results (i), (ii), (iii), and (iv) hold without change.
(7) Even worse is the situation of the solar system, i.e., if one tries to apply the above theorem to Eq. (5.9.2) as a perturbation to Eq. (5.9.1).

The problem lies not so much in the unboundedness of the potentials in the Kepler motions. In fact, in a vicinity W of the Kepler motions of the actual planets, there are no collisions, so the perturbation is bounded there (W has to be thought of as a subset in the nine planets phase space $\mathcal{R}^{27} \times \mathcal{R}^{27}$).

The difficulty lies in the fact that for the unperturbed system described by Eq. (5.9.1), Kepler's laws hold and say that each planet moves periodically with pulsation ω_i; and, therefore, the system moves quasi-periodically with nine independent pulsations instead of the 27 that should be present if the system were really anisochronous and the condition (5.9.5) cannot hold (since two of the three pulsations of each planet i have to be integer multiples of ω_i).

Nevertheless, it is possible to find a version of Proposition 15 covering this problem at least in some nontrivial cases of N gravitating point masses attracted by a fixed center and attracting each other (see also p.493).

Without quoting the exact results, we mention one of their consequences: there exist quasi periodic motions of the planets (i.e., solutions of Eq. (5.9.2)] which take place on almost circular, almost closed, and almost coplanar orbits of distinct radii, provided the masses are very small; hence, there are motions of Eq. (5.9.2) quasi-periodic and without collisions or escapes.

The last statement and result solves a problem which for centuries fascinated physicists, mathematicians, and astronomers. Newton's universal gravitation law is not incompatible, by itself, with the stability of the solar system, a fact empirically observed since millennia and hoped for by everybody. Nevertheless, it remains an open question whether or not our own solar system, modeled by Eq. (5.9.2), is actually stable: the initial data on the positions and velocities of the planets and their masses seem too far from values to which the above-mentioned extensions of Proposition 15 can be applied.
(8) The proof of Proposition 15 gives much more information than its text expresses. It might even be possible to extract from it, and from its extensions mentioned in Observation (6), some astronomically interesting results. However, much work has to be done, since the results of Proposition 15 and its extensions are seldom obtained in "optimal" form. Actually, to my knowledge, careful estimates based on the proof of the theorem and taking full advantage of the peculiarities of a given equation of interest have begun to appear only relatively recently in the simplest cases.

(9) The ideas for the proof of the above theorem arise from perturbation theory for classical Hamiltonian systems: to it the next section is devoted. In §5.11 and §5.12 it will be shown how the ideas of perturbation theory may be applied to prove Proposition 15 in the simplest case of a canonically analytically integrable system, analytically perturbed.

It is a shame that the old classical perturbation theory, which gave rise to analytical mechanics and to the Hamilton-Jacobi method, is nowadays almost forgotten since many people seem to know or care only for the quantum-mechanical perturbation theory. This fact is largely responsible for the aura of mystery which still seems to surround the above theorem.

5.10 Formal Theory of Perturbations. Hamilton–Jacobi Method

Or ti riman, lettor, sovra 'l tuo banco,
Dietro pensando a cio che si preliba,
S'esser vuoi lieto assai prima che stanco.
Messo t'ho innanzi: omai per te ti ciba;
Che a se torce tutta la mia cura
Quella materia ond'io son fatto scriba.[22]

Consider an ℓ-degree-of-freedom system with a Hamiltonian function H on an open set W in phase space. Denote $(\mathbf{p},\mathbf{q})$ the points in W and denote

$$(\mathbf{p},\mathbf{q}) \to H(\mathbf{p},\mathbf{q}), \qquad (\mathbf{p},\mathbf{q}) \in W \tag{5.10.1}$$

the Hamiltonian function H.

We shall suppose that this system is canonically analytically integrable via an analytic canonical transformation I integrating it by transforming W into $V \times \mathcal{T}^\ell$ with $V \subset \mathcal{R}^\ell$, open and bounded.[23]

The transformation I transforms the Hamiltonian H into a function of the first ℓ variables of $(\mathbf{A},\boldsymbol{\varphi}) = I(\mathbf{p},\mathbf{q})$: $h(\mathbf{A}) = H(I^{-1}(\mathbf{A},\boldsymbol{\varphi}))$, $\forall\,(\mathbf{A},\boldsymbol{\varphi}) \in V \times \mathcal{T}^\ell$.

We recall that the variables $\mathbf{A}$ are called "action variables", while the $\boldsymbol{\varphi}$ variables are called "angle variables".

Let F be an analytic function on W and consider the Hamiltonian system described on W by the Hamiltonian function

[22] In basic English:
Now stay, o reader, on your bench,
thinking about what is foreshadowed
if you wish to be happy before being tired.
I did initiate you: now proceed by yourself;
as my whole thoughts are absorbed
by the matter about which I am scribe.
(Dante, Paradiso, Canto X)

[23] See Definition 11, p.289, §4.8.

$$H(\mathbf{p},\mathbf{q}) + \varepsilon F(\mathbf{p},\mathbf{q}) \tag{5.10.2}$$

or, in the action-angle variables, $(\mathbf{A},\boldsymbol{\varphi}) \in V \times \mathcal{T}^\ell$, by

$$h(\mathbf{A}) + \varepsilon f(\mathbf{A},\boldsymbol{\varphi}) \qquad \text{with} \tag{5.10.3}$$

$$h(\mathbf{A}) = H(I^{-1}(\mathbf{A},\boldsymbol{\varphi})), \qquad f(\mathbf{A},\boldsymbol{\varphi}) = F(I^{-1}(\mathbf{A},\boldsymbol{\varphi})) \tag{5.10.4}$$

Perturbation theory proposes to compare, for ε small, the motions of the system with Hamiltonian h and those of the system with Hamiltonian $h+\varepsilon f$, usually with the same initial data.

As stated in §5.9 the comparison methods for solutions of a differential equation depending on a parameter (Lyapunov criterion, attractive manifold theorem, Hopf theorem, etc.) often reveal themselves to be inadequate in the analysis of the problems and difficulties connected with the stability of conservative systems. Such problems appear quite different from those arising in the theory of dissipative systems, at least at the beginning (although the advanced theory ultimately may conceptually coincide).

However, the special form of the Hamiltonian equations permits the use of a simple algorithm, of great interest for applications, for the analysis of the motions of quasi-integrable systems.

The idea is to change variables via a completely canonical transformation $(\mathbf{A},\boldsymbol{\varphi}) \to (\mathbf{A}',\boldsymbol{\varphi}')$, arranging things so that the "old" Hamiltonian (5.10.3) takes the form

$$h_\varepsilon^{(n)}(\mathbf{A}') + \varepsilon^{n+1} f_\varepsilon^{(n)}(\mathbf{A}',\boldsymbol{\varphi}') \tag{5.10.5}$$

in the new variables, where $(\mathbf{A}',\boldsymbol{\varphi}')$ denote the new variables and $h_\varepsilon^{(n)}, f_\varepsilon^{(n)}$ are analytic functions of ε near 0, of $\boldsymbol{\varphi}' \in \mathcal{T}^\ell$ and of $\mathbf{A}'$ in a suitable open set.

Hence, for ε small, the error that would be made supposing that in the variables $(\mathbf{A}',\boldsymbol{\varphi}')$ the system is integrable and described by the Hamiltonian $h_\varepsilon^{(n)}$ is very much smaller than the one that would be made assuming the system as integrable in the original variables $(\mathbf{A},\boldsymbol{\varphi})$ simply setting $\varepsilon = 0$ in Eq. (5.10.3): provided, as we suppose as an extra essential requirement of construction, the canonical transformation itself is not singular at $\varepsilon = 0$.

Intuitively, neglecting in Eq. (5.10.5), or, better, in the Hamiltonian equations associated with Eq. (5.10.5), the $\boldsymbol{\varphi}'$-dependent term produces an error of the order $\varepsilon^{n+1}T$ in the equations solutions, if they are observed up to a time T. Hence, given an approximation η, it will be possible to retain it, although neglecting the influence of f_ε^{n+1} on the motions of Eq. (5.10.5), for a time of the order $T_{\eta,\varepsilon} \propto \eta\varepsilon^{-(n+1)}$.

For ε small this may give substantially better result for $n > 0$ than the one corresponding to the simple, but often too rough, analogous approximation with $n = 0$ (i.e., $\varepsilon = 0$ in Eq. (5.10.3).

The reader will realize that the method that will be used for the "reduction to higher order" of the perturbation via a canonical transformation is nothing

more than a method for constructing successive approximations to the time independent solutions of the Hamilton-Jacobi equation Eq. (3.11.6), p.213.

There are two remarkable cases which can actually be treated along the above lines, building a completely canonical transformation changing Eq. (5.10.3) into Eq. (5.10.5) at least for $(\mathbf{A}, \boldsymbol{\varphi})$ in a neighborhood of the form $S_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell \subset V \times \mathcal{T}^\ell$, where $S_\varrho(\mathbf{A}_0)$ is a sphere with radius ϱ in $\mathcal{R}^\ell$ around a preassigned point $\mathbf{A}_0$, and for some $n > 0$ and ε small.

The first case arises when

$$h(\mathbf{A}) = \boldsymbol{\omega}_0 \cdot (\mathbf{A} - \mathbf{A}_0) \tag{5.10.6}$$

with $\boldsymbol{\omega}_0 \in \mathcal{R}^\ell$ such that there are $C, \alpha > 0$, for which

$$C = \sup_{\boldsymbol{\nu} \in \mathcal{Z}^\ell, \boldsymbol{\nu} \neq \mathbf{0}} \frac{|\boldsymbol{\omega} \cdot \boldsymbol{\nu}|^{-1}}{|\boldsymbol{\nu}|^\alpha} < +\infty \tag{5.10.7}$$

The second case arises when the Fourier coefficients of the development of f:

$$f(\mathbf{A}, \boldsymbol{\varphi}) = \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell} f_{\boldsymbol{\nu}} e^{i\boldsymbol{\nu} \cdot \boldsymbol{\varphi}}, \qquad f_{\boldsymbol{\nu}} = \frac{1}{(2\pi)^\ell} \int_{\mathcal{T}^\ell} f(\mathbf{A}, \boldsymbol{\varphi}) e^{-i\boldsymbol{\nu} \cdot \boldsymbol{\varphi}} \, d\boldsymbol{\varphi} \tag{5.10.8}$$

vanish for $|\boldsymbol{\nu}| > N$ and, setting

$$\boldsymbol{\omega}(\mathbf{A}) = \frac{\partial h}{\partial \mathbf{A}}, \qquad \text{one has} \tag{5.10.9}$$

$$|\boldsymbol{\omega}(\mathbf{A}_0) \cdot \boldsymbol{\nu}| > 0, \quad \forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell,\ 0 < |\boldsymbol{\nu}| \le N. \tag{5.10.10}$$

In the first case, it is even possible to put the Hamiltonian into the form of Eq. (5.10.5), $\forall\, n = 0, 1, \ldots$, provided ε is small enough (depending, however, on the choice of n).

The above statements are illustrated in the following classical propositions.

16 Proposition. *Consider the Hamiltonian (5.10.3) on $V \times \mathcal{T}^\ell$ with*

$$f(\mathbf{A}, \boldsymbol{\varphi}) = \sum_{\substack{\boldsymbol{\nu} \in \mathcal{Z}^\ell \\ |\boldsymbol{\nu}| \le N}} f_{\boldsymbol{\nu}} e^{i\boldsymbol{\nu} \cdot \boldsymbol{\varphi}} \tag{5.10.11}$$

analytic on $V \times \mathcal{T}^\ell$, with $N > 0$, and suppose that, $\forall\, \mathbf{A}_0 \in V$, the function h is such that

$$|\boldsymbol{\omega}(\mathbf{A}_0) \cdot \boldsymbol{\nu}| > 0, \quad \forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell,\ 0 < |\boldsymbol{\nu}| \le N. \tag{5.10.12}$$

Then there exist $\varrho_1 > 0, \varepsilon_1 > 0$ and, $\forall\, \varepsilon \in (-\varepsilon_1, \varepsilon_1)$, a completely canonical transformation $(\mathbf{A}, \boldsymbol{\varphi}) \longleftrightarrow (\mathbf{A}', \boldsymbol{\varphi}')$ defined for $(\mathbf{A}, \boldsymbol{\varphi}) \in W_\varepsilon$, with $V \times \mathcal{T}^\ell \supset W_\varepsilon \supset S_{\frac{1}{2}\varrho_1}(\mathbf{A}_0) \times \mathcal{T}^\ell$ and with values onto $S_{\varrho_1}(\mathbf{A}_0) \times \mathcal{T}^\ell$, smoothly depending on ε and transforming the Hamiltonian (5.10.3) into

$$h_\varepsilon^{(1)}(\mathbf{A}') + \varepsilon^2 f_\varepsilon^{(1)}(\mathbf{A}',\boldsymbol{\varphi}), \tag{5.10.13}$$

where $h_\varepsilon^{(1)}, f_\varepsilon^{(1)}$ *are analytic in* $\varepsilon, \mathbf{A}', \boldsymbol{\varphi}'$. *Furthermore* $h_\varepsilon^{(1)}$ *can be given a simple expression; see Eq. (5.10.25) below.*

Observation. As mentioned above, the reader should interpret the proof that follows as a "perturbative solution to order ε" of the Hamilton-Jacobi equation in the time-independent case, i.e., when H in Eq. (3.11.68), p.226, does not explicitly depend on t. Actually, the above proposition is the basic example of how the method of Hamilton-Jacobi concretely works. Most applications of the Hamilton-Jacobi's method are based on this proposition.

PROOF. The canonical transformation will be determined by looking for a generating function Φ, see §3.11 and §3.12 from p.222 on.

Such a transformation is expected to be close to the identity up to infinitesimals $O(\varepsilon)$, thus the unknown generating function will be written as

$$\mathbf{A}' \cdot \boldsymbol{\varphi} + \Phi(\mathbf{A}', \boldsymbol{\varphi}), \tag{5.10.14}$$

where $(\mathbf{A}', \boldsymbol{\varphi}) \to \mathbf{A}' \cdot \boldsymbol{\varphi}$ is the generating function of the identity map and Φ is infinitesimal in ε . The function Φ will be determined by requiring that the Hamiltonian in the new variables $(\mathbf{A}', \boldsymbol{\varphi}')$ defined by the formal map

$$\begin{aligned} \mathbf{A} =& \mathbf{A}' + \frac{\partial \Phi}{\partial \boldsymbol{\varphi}}(\mathbf{A}', \boldsymbol{\varphi}), \\ \boldsymbol{\varphi}' =& \boldsymbol{\varphi} + \frac{\partial \Phi}{\partial \mathbf{A}'}(\mathbf{A}', \boldsymbol{\varphi}), \end{aligned} \tag{5.10.15}$$

i.e., the function

$$h\big(\mathbf{A}' + \frac{\partial \Phi}{\partial \boldsymbol{\varphi}}(\mathbf{A}', \boldsymbol{\varphi})\big) + \varepsilon f\big(\mathbf{A}' + \frac{\partial \Phi}{\partial \boldsymbol{\varphi}}(\mathbf{A}', \boldsymbol{\varphi}), \boldsymbol{\varphi}\big) \tag{5.10.16}$$

is $\boldsymbol{\varphi}$ independent up to terms infinitesimal of higher order in ε.

Since, as already said, we expect that $\Phi \simeq O(\varepsilon)$, we can heuristically find, by developing Eq. (5.10.16) in series with respect to $\frac{\partial \Phi}{\partial \mathbf{A}'}$, that the equation for Φ (the "Hamilton-Jacobi equation to first order in ε") is

$$\frac{\partial h}{\partial \mathbf{A}'}(\mathbf{A}') \cdot \frac{\partial \Phi}{\partial \boldsymbol{\varphi}}(\mathbf{A}', \boldsymbol{\varphi}) + \varepsilon f(\mathbf{A}', \boldsymbol{\varphi}) = \{\boldsymbol{\varphi} - \text{independent function}\} \tag{5.10.17}$$

which, written in terms of the Fourier components of Φ, means that if

$$i\,(\boldsymbol{\omega}(\mathbf{A}') \cdot \boldsymbol{\nu})\Phi_{\boldsymbol{\nu}}(\mathbf{A}') + \varepsilon\, f_{\boldsymbol{\nu}}(\mathbf{A}') = 0, \quad \forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell,\ |\boldsymbol{\nu}| > 0 \tag{5.10.18}$$

This equation is really a soluble equation if $|\mathbf{A}' - \mathbf{A}_0| \le \overline{\varrho}_1$, with $\overline{\varrho}_1$, so small that the closure of $S_{\overline{\varrho}_1}(\mathbf{A}_0)$ is a subset of V and therefore

$$\boldsymbol{\omega}(\mathbf{A}')\cdot\boldsymbol{\nu}\neq 0,\quad \forall\,\boldsymbol{\nu}\in\mathcal{Z}^\ell,\ 0<|\boldsymbol{\nu}|\le N. \tag{5.10.19}$$

see Eq. (5.10.12). Then we can define in $S_{\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ the analytic function

$$\Phi(\mathbf{A}',\boldsymbol{\varphi})=\varepsilon\sum_{0<|\boldsymbol{\nu}|\le N}\frac{f_{\boldsymbol{\nu}}(\mathbf{A}')e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i\boldsymbol{\omega}(\mathbf{A}')\cdot\boldsymbol{\nu}}. \tag{5.10.20}$$

It follows from the implicit function theorem, see Appendix G, Corollaries 3 and 4, that the second of Eqs. (5.10.15) can be uniquely inverted with respect to $\boldsymbol{\varphi}$ and the first of Eqs. (5.10.15) can be inverted with respect to $\mathbf{A}'$ in the respective forms

$$\begin{aligned}\boldsymbol{\varphi}=&\boldsymbol{\varphi}'+\boldsymbol{\Delta}(\mathbf{A}',\boldsymbol{\varphi}'), &\qquad \boldsymbol{\Delta}\in\mathcal{C}^\infty(\overline{S_{\varrho_1}(\mathbf{A}_0)}\times\mathcal{T}^\ell)\\ \mathbf{A}'=&\mathbf{A}+\boldsymbol{\Xi}'(\mathbf{A},\boldsymbol{\varphi}), &\qquad \boldsymbol{\Xi}\in\mathcal{C}^\infty(\overline{S_{\varrho_1}(\mathbf{A}_0)}\times\mathcal{T}^\ell)\end{aligned} \tag{5.10.21}$$

if ε is small enough,[24] i.e., if $|\varepsilon|<\widetilde{\varepsilon}_1$, with $\widetilde{\varepsilon}_1$, suitably chosen; and also there is $B>0$ such that

$$|\frac{\partial\Phi}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi})|=|\boldsymbol{\Xi}'(\mathbf{A},\boldsymbol{\varphi})|<B\,|\varepsilon|, \tag{5.10.22}$$

so that, if $B|\varepsilon|<\frac{1}{8}\overline{\varrho}_1$, the maps $(\mathbf{A}',\boldsymbol{\varphi}')\to\mathcal{C}(\mathbf{A}',\boldsymbol{\varphi}')=(\mathbf{A},\boldsymbol{\varphi})$:

$$\begin{aligned}\mathbf{A}=&\mathbf{A}'+\frac{\partial\Phi}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi}'+\boldsymbol{\Delta}(\mathbf{A}',\boldsymbol{\varphi}')),\\ \boldsymbol{\varphi}=&\boldsymbol{\varphi}'+\boldsymbol{\Delta}(\mathbf{A}',\boldsymbol{\varphi}')\end{aligned} \tag{5.10.23}$$

and $(\mathbf{A},\boldsymbol{\varphi})\to\mathcal{C}'(\mathbf{A},\boldsymbol{\varphi})=(\mathbf{A}',\boldsymbol{\varphi}')$:

$$\begin{aligned}\mathbf{A}'=&\mathbf{A}+\boldsymbol{\Xi}'(\mathbf{A},\boldsymbol{\varphi}),\\ \boldsymbol{\varphi}'=&\boldsymbol{\varphi}+\frac{\partial\Phi}{\partial\mathbf{A}'}(\mathbf{A}+\boldsymbol{\Xi}'(\mathbf{A},\boldsymbol{\varphi}),\boldsymbol{\varphi})\end{aligned} \tag{5.10.24}$$

are well defined on $S_{\frac{1}{2}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ and take values in $S_{\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$. Furthermore, $\mathcal{C}$ and $\mathcal{C}'$ map $S_{\frac{1}{4}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ into $S_{\frac{1}{2}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ and $\mathcal{C}\mathcal{C}'=\mathcal{C}'\mathcal{C}=\{$ identity map$\}$ on $S_{\frac{1}{4}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ by construction (and by the uniqueness part of the implicit function theorem).

Therefore, the Jacobian determinants of $\mathcal{C}$ or $\mathcal{C}'$' on $S_{\frac{1}{4}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ cannot vanish and, hence, by Proposition 21, §3.11, p.220, $\mathcal{C}$ is a completely canonical map of $S_{\frac{1}{4}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$ onto its image $W_\varepsilon\supset S_{\frac{1}{8}\overline{\varrho}_1}(\mathbf{A}_0)\times\mathcal{T}^\ell$. So we take $\varepsilon_1=\min(\widetilde{\varepsilon}_1,\frac{1}{8}\overline{\varrho}_1)$.

By the construction of Φ [see Eq. (5.10.17)] the Hamiltonian function in the $(\mathbf{A}',\boldsymbol{\varphi}')$ variables has the form of Eq. (5.10.13). By substituting Eq. (5.10.20) into Eq. (5.10.17), one, in fact, also obtains

[24] $\boldsymbol{\Xi}'$ and $\boldsymbol{\Delta}$ are C^∞ also in ε, jointly with $(\mathbf{A},\boldsymbol{\varphi})$ or $(\mathbf{A}',\boldsymbol{\varphi}')$, by the implicit functions theorems in Appendix G.

$$h_\varepsilon^{(1)}(\mathbf{A}') = h(\mathbf{A}') + \varepsilon f_{\mathbf{0}}(\mathbf{A}'), \tag{5.10.25}$$

where $f_{\mathbf{0}}$ is the **0**-th Fourier coefficient of f, see Eq. (5.10.8).

The analyticity of the canonical maps $\mathcal{C}$ and $\mathcal{C}'$ will not be discussed here. It follows if Eqs. (5.10.21) are obtained via the application of analytic implicit function theorems that will be discussed in the next section; see Propositions 18-20. mbe

The above discussion is the basis for the most common algorithms in the calculations of the perturbed Hamiltonian motions; it leads to the natural idea of iterating the procedure by reducing the perturbation from $O(\varepsilon^2)$ to $O(\varepsilon^4)$, etc.

The difficulty lies in the fact that, in general, the new Hamiltonian (5.10.16) which, to first order in ε reduces to Eq. (5.10.25), no longer has the form necessary for applicability of Proposition 16. In fact, the perturbation of order ε^2 will be a function of $(\mathbf{A}', \boldsymbol{\varphi}')$ which has *all, or at least infinitely many*, harmonic components in $\boldsymbol{\varphi}'$ non vanishing, disregarding exceptional cases. One can convince oneself of this with some thought, noting that $\frac{\partial\Phi}{\partial\boldsymbol{\varphi}}(\mathbf{A}', \boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}', \boldsymbol{\varphi}'))$ contains terms like $e^{i\boldsymbol{\Delta}(\mathbf{A}', \boldsymbol{\varphi}')\cdot\boldsymbol{\nu}}$ and, unless some "miraculous" cancellations take place, will no longer be trigonometric polynomials in $\boldsymbol{\varphi}'$.

The following proposition, valid in the other case considered in the introduction to Proposition 16, is quite interesting because it shows that with a slight modification of the method of the above proof but under different assumptions, one can "remove" the perturbation to an arbitrary order in ε.

17 Proposition. *Consider the Hamiltonian function given by Eq. (5.10.3) on $V \times \mathcal{T}^\ell$ with h verifying Eqs. (5.10.6) and (5.10.7) and f analytic. There is $\varrho > 0$ such that:*
(1) For each $n = 0, 1 \ldots$. there exists $\varepsilon_n > 0$ and, $\forall\, |\varepsilon| < \varepsilon_n$, functions $\Phi_{\varepsilon,n}$ defined on $S_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ and analytic in ε and in the other arguments $(\mathbf{A}, \boldsymbol{\varphi})$, generating completely canonical transformation $(\mathbf{A}, \boldsymbol{\varphi}) \longleftrightarrow (\mathbf{A}', \boldsymbol{\varphi}')$ such that

$$\begin{aligned} \mathbf{A} &= \mathbf{A}' + \frac{\partial \Phi_{\varepsilon,n}}{\partial \boldsymbol{\varphi}}(\mathbf{A}', \boldsymbol{\varphi}), \\ \boldsymbol{\varphi}' &= \boldsymbol{\varphi} + \frac{\partial \Phi_{\varepsilon,n}}{\partial \mathbf{A}'}(\mathbf{A}', \boldsymbol{\varphi}), \end{aligned} \tag{5.10.26}$$

mapping a subset $W_{\varepsilon,n}$, $S_{\frac{1}{2}\varrho}(\mathbf{A}_0) \times \mathcal{T}^\ell \subset W_{\varepsilon,n} \subset V \times \mathcal{T}^\ell$, onto $S_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$.
(2) The map of Eq. (5.10.26) transforms the Hamiltonian into the form ("Birkhoff normal form")

$$h_{\varepsilon,n}(\mathbf{A}') + \varepsilon^{n+1} f_\varepsilon^{(n)}(\mathbf{A}', \boldsymbol{\varphi}') \tag{5.10.27}$$

where $h_{\varepsilon,n}(\mathbf{A}')$ is analytic in $\varepsilon, \mathbf{A}'$ and $f_\varepsilon^{(n)}$ is also analytic in $\varepsilon, \mathbf{A}', \boldsymbol{\varphi}'$. An explicit expression for $h_{\varepsilon,n}(\mathbf{A}')$ is Eq. (5.10.41).

Observation. The construction described in the proof of this proposition is often referred to as the "Birkhoff transformation".

PROOF. Define heuristically:

$$\Phi_{\varepsilon,n}(\mathbf{A}',\boldsymbol{\varphi}) = \sum_{k=1}^{n} \varepsilon^k \Phi^{(k)}(\mathbf{A}',\boldsymbol{\varphi}) \tag{5.10.28}$$

and consider the Hamiltonian in the new variables $(\mathbf{A}',\boldsymbol{\varphi})$, Eq. (5.10.26):

$$h\big(\mathbf{A}' + \frac{\partial \Phi_{\varepsilon,n}}{\partial \boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi})\big) + \varepsilon f\big(\mathbf{A}' + \frac{\partial \Phi_{\varepsilon,n}}{\partial \boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi}),\boldsymbol{\varphi}\big). \tag{5.10.29}$$

Developing this expression in powers of ε using the analyticity of f and h (the latter is actually linear) in $\mathbf{A}$, impose that the resulting series in ε,

$$\sum_{k=1}^{\infty} \psi^{(k)}(\mathbf{A}',\boldsymbol{\varphi})\,\varepsilon^k, \tag{5.10.30}$$

has all the coefficients $\psi^{(k)}$, $k = 0, 1, \ldots, n$, $\boldsymbol{\varphi}$-independent.

This condition allows one to determine recursively $\Phi^{(1)}, \ldots, \Phi^{(n)}$ [and it appears that $\varepsilon\Phi^{(1)}$ is given by Eq. (5.10.20), of course].

Then, once the expressions for $\Phi^{(1)}, \ldots, \Phi^{(n)}$ are found, one shall write Eq. (5.10.26), and by taking ε small, proceeding exactly as in the proof of Proposition 16, the implicit function theorem will be used to guarantee that Eq. (5.10.26) actually defines a canonical transformation between $S_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ and some $W_{\varepsilon,n} \subset V \times \mathcal{T}^\ell$ and $W_{\varepsilon,n} \supset S_{\frac{1}{2}\varrho}(\mathbf{A}_0) \times \mathcal{T}^\ell$. The invertibility conditions will depend on n. By construction, Eq. (5.10.27) will then follow, with $h_{\varepsilon,n}, f_\varepsilon^{(n)}$ of class C^∞ in $\varepsilon, \mathbf{A}', \boldsymbol{\varphi}'$. They are actually analytic and this point can be commented as at the end of the proof of Proposition 16, see p.469.

Hence, the whole problem is to show that one can find $\Phi^{(1)}, \ldots, \Phi^{(n)}$ so that the formal series of Eq. (5.10.30) has the first $(n+1)$ coefficients with harmonics in $\boldsymbol{\varphi}$ of order $\boldsymbol{\nu} \neq \mathbf{0}$ vanishing. This is a purely algebraic problem.

As amply exploited in the following section, where the question will be more systematically treated, the analyticity assumption on f implies that it can be developed in the Taylor series about $\mathbf{A}_0$ and in the Fourier series in $\boldsymbol{\varphi}$ in the form

$$f(\mathbf{A},\boldsymbol{\varphi}) = \sum_{\mathbf{a}\mathcal{Z}_+^\ell,\boldsymbol{\nu}\in\mathcal{Z}^\ell} f_{\boldsymbol{\nu}}^{(\mathbf{a})}(\mathbf{A}-\mathbf{A}_0)^{\mathbf{a}} e^{i\,\boldsymbol{\nu}\cdot\boldsymbol{\varphi}} = \sum_{\mathbf{a}\in\mathcal{Z}_+^\ell} f^{(\mathbf{a})}(\mathbf{A}_0,\boldsymbol{\varphi})(\mathbf{A}-\mathbf{A}_0)^{\mathbf{a}} \tag{5.10.31}$$

where, see Definition 13, p.336, $\mathbf{a} = (\alpha_1, \ldots, a_\ell) \in \mathcal{Z}_+^\ell$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_\ell) \in \mathcal{Z}^\ell$ and

$$(\mathbf{A}-\mathbf{A}_0)^{\mathbf{a}} = \prod_{i=1}^{\ell}(A_i - A_{0i})^{a_i}, \qquad \boldsymbol{\nu}\cdot\boldsymbol{\varphi} = \sum_{i=1}^{\ell}\nu_i\varphi_i \tag{5.10.32}$$

and, furthermore, there are $R>0, \varrho_0>0, \xi_0>0$, such that

$$|f_{\boldsymbol{\nu}}^{(\mathbf{a})}| \le R\,\varrho_0^{-|\mathbf{a}|}e^{-\xi_0|\boldsymbol{\nu}|}, \qquad \forall\ \mathbf{a}\in\mathcal{Z}_+^{\ell},\ \forall\ \boldsymbol{\nu}\in\mathcal{Z}^{\ell}, \tag{5.10.33}$$

if $|\mathbf{a}| \stackrel{def}{=} \sum_{i=1}^{\ell} a_i$, $|\boldsymbol{\nu}| \stackrel{def}{=} \sum_{i=1}^{\ell}|\nu_i|$.

This inequality is not immediately obvious and it will be discussed in §5.11; for the time being, we suppose and use Eq. (5.10.33) without discussion.

Developing Eq. (5.10.29) in powers of ε and collecting the terms of equal order in ε and setting $f^{(\mathbf{a})}(\mathbf{A}',\boldsymbol{\varphi}) = \frac{1}{\mathbf{a}!}\frac{\partial^{|\mathbf{a}|}f(\mathbf{A}',\boldsymbol{\varphi})}{(\partial\mathbf{A}')^{\mathbf{a}}}$ with $\mathbf{a}! \stackrel{def}{=} \prod_{i=1}^{\ell} a_i!$ (it is the $\mathbf{a}$-th coefficient of the Taylor expansion of f around $\mathbf{A}'$ at fixed $\boldsymbol{\varphi}$), one finds [using Eq. (5.10.31)]

$$\psi^{(k)}(\mathbf{A}',\boldsymbol{\varphi}) = \Big\{\sum_{\mathbf{a}\in\mathcal{Z}_+^{\ell}} f^{(\mathbf{a})}(\mathbf{A}',\boldsymbol{\varphi}) \sum^{*}_{\mathbf{n}^1_{a_1},\ldots,\mathbf{n}^{\ell}_{a_\ell}} \prod_{j=1}^{\ell}\Big(\prod_{s=1}^{a_j}\frac{\partial\Phi^{(n_s^j)}(\mathbf{A}',\boldsymbol{\varphi})}{\partial\varphi_j}\Big)\Big\}$$

$$+\boldsymbol{\omega}_0\cdot\frac{\partial\Phi^{(k)}}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi}) \stackrel{def}{=} \{N^{(k)}(\mathbf{A}',\boldsymbol{\varphi})\} + \boldsymbol{\omega}_0\cdot\frac{\partial\Phi^{(k)}}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi}) \tag{5.10.34}$$

for $k=1,2,\ldots$, and the $*$ means that the sum is performed subject to the constraint $\sum_{j=1}^{\ell}\sum_{s=1}^{a_j} n_s^j = k-1$. Furthermore, we set

$$\psi^(0)(\mathbf{A}',\boldsymbol{\varphi}) = h(\mathbf{A}') \tag{5.10.35}$$

The condition that $\psi^{(1)}$ is $\boldsymbol{\varphi}$-independent (hence, $\boldsymbol{\varphi}'$ independent) becomes, by Eq. (5.10.34),

$$f(\mathbf{A}',\boldsymbol{\varphi}) + \boldsymbol{\omega}_0\cdot\frac{\partial\Phi^{(1)}(\mathbf{A}',\boldsymbol{\varphi})}{\partial\boldsymbol{\varphi}} = \{\boldsymbol{\varphi} - \text{independent function}\} \tag{5.10.36}$$

and it determines $\Phi^{(1)}$, up to a function of $\mathbf{A}'$ alone, as:

$$\Phi^{(1)}(\mathbf{A}',\boldsymbol{\varphi}) = \sum_{\mathbf{a}\in\mathcal{Z}_+^{\ell}}\frac{f_{\boldsymbol{\nu}}^{(\mathbf{a})}(\mathbf{A}')e^{i\,\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i\,\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}}, \tag{5.10.37}$$

where $f_{\boldsymbol{\nu}}(\mathbf{A}')$ is the $\boldsymbol{\nu}$-th Fourier coefficient of $f(\mathbf{A}',\boldsymbol{\varphi})$ at $\mathbf{A}'$ fixed:

$$f_{\boldsymbol{\nu}}(\mathbf{A}') = \sum_{\mathbf{a}\in\mathcal{Z}_+^{\ell}} f^{(\mathbf{a})}(\mathbf{A}_0)(\mathbf{A}'-\mathbf{A}_0)^{\mathbf{a}}. \tag{5.10.38}$$

Replacing $f_{\boldsymbol{\nu}}(\mathbf{A}')$ in Eq. (5.10.37) by Eq. (5.10.38) and using Eqs. (5.10.33) and (5.10.7), one sees that the series in Eq. (5.10.37) converges and defines a C^{∞} function of $(\mathbf{A}',\boldsymbol{\varphi})\in S_{\varrho_0}(\mathbf{A}_0)\times\mathcal{T}^{\ell}$ (actually such a function is analytic, as could be shown).

Then, from Eq. (5.10.34), it follows that

$$N^{(2)}(\mathbf{A}',\boldsymbol{\varphi}) = \sum_{j=1}^{\ell} f^{(\mathbf{e}_j)}(\mathbf{A}',\boldsymbol{\varphi}) \frac{\partial \Phi^{(1)}(\mathbf{A}',\boldsymbol{\varphi})}{\partial \varphi_j} \tag{5.10.39}$$

with $\mathbf{e}_1 = (1,0,\ldots,0)$, $\mathbf{e}_2 = (0,1,\ldots,0)\ldots$.

From what has been said above, it follows that $N^{(2)}$ is a $C^\infty(S_{\varrho_0}(\mathbf{A}_0) \times \mathcal{T}^\ell)$ function (actually analytic), and if $N^{(2)}_{\boldsymbol{\nu}}(\mathbf{A}')$ denotes its $\boldsymbol{\nu}$-th Fourier coefficient, the condition that $\psi^{(2)}$ in Eq. (5.10.34) is $\boldsymbol{\varphi}$-independent yields

$$\Phi^{(2)}(\mathbf{A}',\boldsymbol{\varphi}) = \sum_{\mathbf{0}\neq\boldsymbol{\nu}\in\mathcal{Z}^\ell} \frac{N^{(2)}(\mathbf{A}')e^{i\,\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i\,\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}}, \tag{5.10.40}$$

which, again from Eq. (5.10.33) and from Eqs. (5.10.31), (5.10.37), and (5.10.38), turns out to be a C^∞ function on $S_{\varrho_0}(\mathbf{A}_0)\times\mathcal{T}^\ell$ (actually analytic), etc., inductively. Hence

$$h_{\varepsilon,n}(\mathbf{A}') = h_0(\mathbf{A}') + \sum_{k=1}^{n} \varepsilon^k\, N^{(k)}_{\mathbf{0}}(\mathbf{A}'). \tag{5.10.41}$$

mbe

Observations.
(1) Equations (5.10.37) and (5.10.40) and their generalizations to higher k show that $N^{(k)}(\mathbf{A}',\boldsymbol{\varphi})$ can be chosen to be n independent. It becomes natural to consider the limit as $n\to\infty$. In this limit, the perturbation would disappear and the Hamiltonian would be transformed into

$$h_\varepsilon(\mathbf{A}') = h(\mathbf{A}') + \sum_{k=1}^{\infty} \varepsilon^k\, N^{(k)}_{\mathbf{0}}(\mathbf{A}').$$

and it would therefore be integrable. However, the estimates on ε_n that can be derived by applying the scheme suggested in the above proof appear to be such that $\varepsilon_n \xrightarrow[n\to+\infty]{} 0$, save some exceptional cases. Therefore, nothing can be concluded about the limit $n\to+\infty$.
It is known that it cannot happen, in general, that both series ("Birkhoff's formal series").

$$\sum_{k=1}^{\infty} \varepsilon^k\, N^{(k)}_{\mathbf{0}}(\mathbf{A}'), \qquad \sum_{k=1}^{\infty} \varepsilon^k\, \Phi^{(k)}(\mathbf{A}',\boldsymbol{\varphi}) \tag{5.10.42}$$

converge, defining analytic functions of $(\mathbf{A}',\boldsymbol{\varphi},\varepsilon)$ in $(\mathbf{A}',\boldsymbol{\varphi})\in S_{\varrho_0}(\mathbf{A}_0)\times\mathcal{T}^\ell$ and in ε near zero and, at the same time, $\varepsilon^{n+1}\partial f^{(n)}_\varepsilon \xrightarrow[n\to+\infty]{} 0$ uniformly in the same region of $(\mathbf{A}',\boldsymbol{\varphi},\varepsilon)$.
This would, in fact, imply the existence of ℓ prime integrals analytic in $\varepsilon,\mathbf{A},\boldsymbol{\varphi}$ for ε close to 0, $\mathbf{A}$ close to $\mathbf{A}_0$ and $\boldsymbol{\varphi}\in\mathcal{T}^\ell$: namely, $(A'_1,\ldots,A'_\ell)$, and via such

integrals ("uniform integrals"), the system would be analytically integrable, with a canonical transformation with such integrals as the new action variables. This property has been shown to be impossible in a number of interesting cases.
A simple example in which the series (5.10.42) can be explicitly computed is in Problem 16 at the end of this section: in the example the second of (5.10.42) does not converge. However if $N^{(k)}(\mathbf{A})$ depend on $\mathbf{A}$ via $\boldsymbol{\omega}\cdot\mathbf{A}$ only, then the series converge: this is a nice criterion (see [44]).
(2) Various algorithms used in practice to study perturbations of integrable motions are based on the two propositions illustrated above. The simplest is the following.

First, develop f in a Fourier series. This usually causes great problems. In fact, it is often possible to compute only a few Fourier coefficients for f. However, on the other hand, such coefficients often decrease, as $\boldsymbol{\nu}\to\infty$, very quickly. Then, if f is written as

$$f = f^{[\le N]} + f^{[>N]}, \tag{5.10.43}$$

where, for f given by Eq. (5.10.31), we set [see Eq. (5.10.38)]

$$f^{[\le N]}(\mathbf{A},\boldsymbol{\varphi}) = \sum_{\substack{\mathbf{a}\in\mathcal{Z}^\ell_+,\boldsymbol{\nu}\in\mathcal{Z}^\ell\\ |\boldsymbol{\nu}|\le N}} f^{(\mathbf{a})}_{\boldsymbol{\nu}}(\mathbf{A}-\mathbf{A}_0)^{\mathbf{a}}e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}} \equiv \sum_{\substack{\boldsymbol{\nu}\in\mathcal{Z}^\ell\\ |\boldsymbol{\nu}|\le N}} f_\nu(\mathbf{A})e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}} \tag{5.10.44}$$

one has that $\varepsilon f^{[>N]}$ is very small even for N not too large and its contribution to the Hamiltonian equation produces an error, in a fixed given time, much smaller than $O(\varepsilon)$, say $O(\varepsilon\eta)$ with $\eta\ll 1$.
It is then possible to apply Proposition 16 to the system with Hamiltonian $h+\varepsilon f^{[>N]}$ and remove the perturbation to $O(\varepsilon^2)$. In the new variables, neglecting the perturbation of $O(\varepsilon^2)$ will cause an error, over a fixed time, of order $O(\varepsilon^2+\varepsilon\eta)$ on the solutions of the original equations. This is often a very good approximation if $\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}\ne 0$, $\forall\, 0<|\boldsymbol{\nu}|\le N$, $\forall\,\mathbf{A}\in$ {set of interesting initial actions}.
(3) A special case of great importance to which, however, the above algorithm cannot be applied directly is that of the perturbations of the motion of the Kepler system when, in defining the unperturbed system, one neglects the reciprocal attractions between the planets [i.e., one takes Eq. (5.9.2) as a perturbation of Eq. (5.9.1)].

As we saw, the Kepler motions are rigorously periodic, and to every planet a single pulsation is associated rather than three: the other two vanish (or are integer multiples of the first, depending on which variables are chosen to integrate the motion) as a consequence of the conservation of angular momentum and of the wonderful nature of the Newtonian force which singles it out among the central forces as the most impressive, see §4.9 and §4.10.

It is therefore certainly impossible to satisfy Eq. (5.10.10) with reasonable N. Hence, the above approximation scheme cannot be applied.

Nevertheless, a similar scheme can be applied. Consider the motions in action-angle coordinates $(\mathbf{A}, \boldsymbol{\varphi})$, where $\mathbf{A} = (\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(n)})$, $\boldsymbol{\varphi} = (\boldsymbol{\varphi}^{(1)}, \ldots, \boldsymbol{\varphi}^{(n)})$, where $(\mathbf{A}^{(j)}, \boldsymbol{\varphi}^{(j)})$ are the natural variables, for the systems Sun-i-th planet, in terms of which the Hamiltonian takes the form (if $M_s =$ Sun mass), $m_i = i$-th planet mass, $\varepsilon_{ij} = \frac{\sqrt{m_i m_j}}{M_s}$, see p.458:

$$h_0(\mathbf{A}) - \sum_{i<j} \frac{\varepsilon\, m_i m_j}{|\mathbf{x}_i - \mathbf{x}_j|} \equiv h_0(\mathbf{A}) - \sum_{i<j} \varepsilon_{ij} \frac{K\sqrt{m_i m_j}}{|\mathbf{x}_i - \mathbf{x}_j|}, \tag{5.10.45}$$

$$h_0(\mathbf{A}) = \sum_{i=1}^{n} \overline{h}_0(A_1^{(i)}), \tag{5.10.46}$$

having denoted $A_1^{(j)}$ the first component of $\mathbf{A}^{(j)} = (A_1^{(j)}, A_2^{(j)}, A_3^{(j)})$, and we recall that $\mathbf{A}^{(j)}$ can be chosen as follows (see problems for §4.10):

$$\begin{aligned} A_1^{(j)} =& m_j \omega_j a_j^2 = \frac{(\varepsilon M_s)^{\frac{3}{2}} m_j}{(-2E_j)^{\frac{1}{2}}} \stackrel{def}{=} L_j, \\ A_2^{(j)} =& m_j A(j) \stackrel{def}{=} G_j, \\ A_3^{(j)} =& m_i A(j) \cos i(j) \stackrel{def}{=} \Theta_j, \end{aligned} \tag{5.10.47}$$

where $\frac{1}{2}A(j)$ is the areal velocity of the j-th planet, E_j its energy, a_j is the major semiaxis of its orbit, and $i(j)$ is the inclination of the jth orbit on the ecliptic plane (the ecliptic plane is traditionally the plane of the Earth orbit or more precisely a reference plane fixed with the stars and parallel to a conventional average plane of the Earth orbit).

The angle variables associated with such action variables are $\varphi_1^{(j)} = \ell^{(j)}, \varphi_2^{(j)} = g^{(j)}, \varphi_3^{(j)} = h^{(j)}$ known in astronomy as the "average anomaly", the "major semiaxis longitude" and the "node-line longitude" with respect to the fixed axes established on the ecliptic plane (i.e., on the xy plane of the chosen inertial frame); see Problems 11 and following to §4.10, p.303, for a discussion of these variables.

Equation (5.10.46) shows that in the unperturbed motions, $g^{(j)}, h^{(j)}$ are constants (i.e., $\omega_2^{(j)} = \omega_3^{(j)} = 0$) since h_0 only depends on the variables L_j, $j = 1, \ldots, n$.

One can then proceed to write the perturbation in Eq. (5.10.45) in terms of the $(\mathbf{A}, \boldsymbol{\varphi})$ coordinates (a nontrivial task, in practice; see Problem 15, p.305, §4.10 for the similar question in the case of the planar problem), and afterwards one can try to apply the scheme seen in the proof of Proposition 16 to build a canonical map $(\mathbf{A}, \boldsymbol{\varphi}) \to (\mathbf{A}', \boldsymbol{\varphi}')$ transforming Eq. (5.10.45) into a function independent on the $\varphi_1^{(j)}$ variables, $j = 1, \ldots, n$, to first order in $\overline{\varepsilon} = \max \varepsilon_{ij}$. One shall proceed as prescribed in the proof of Proposition 16, considering $\varphi_2^{(j)}, \varphi_3^{(j)}$ as parameters.

If we call $f(\mathbf{A}, \boldsymbol{\varphi})$ the perturbation term of Eq. (5.10.45), when expressed in the action-angle variables apt to describe the unperturbed system, we introduce the new canonical variables $(\mathbf{A}', \boldsymbol{\varphi}')$ via the generating function $\mathbf{A}' \cdot \boldsymbol{\varphi} + \Phi(\mathbf{A}', \boldsymbol{\varphi})$ with

$$\Phi(\mathbf{A}', \boldsymbol{\varphi}) = \sum_{\substack{\boldsymbol{\nu} \in \mathcal{Z}^n \\ 0<|\boldsymbol{\nu}|\leq N}} \frac{f_{\boldsymbol{\nu}}(\mathbf{A}') e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i \sum_{j=1}^n \nu_1^{(j)} \omega_j(A_j^{'(1)})}, \tag{5.10.48}$$

where N is a "large number" which we imagine here to have chosen such that for some a priori given purposes, neglecting $f^{[>N]}$ in Eq. (5.10.45), produces a negligible error. In this way we obtain a Hamiltonian having the form

$$\begin{aligned} &h(A_1^{'(1)}, \ldots, \varphi_1^{'(1)}, \ldots) = h_0(\mathbf{A}') \\ &+ \varepsilon h_1(A_1^{'(1)}, \ldots, A_3^{'(n)}; \varphi_2^{'(2)}, \ldots \varphi_n^{'(3)}) + O(\varepsilon^2), \end{aligned} \tag{5.10.49}$$

and the equations of motion will become, $j = 1, \ldots, n$,

$$\begin{aligned} \dot{A}_1^{(j)} =&0, \qquad j = 1, \ldots, n \\ \dot{\varphi}_\sigma^{'(j)} =&\varepsilon \frac{\partial h_1}{\partial A_\sigma^{(j)}} (A_1^{'(1)}, \ldots, A_3^{'(n)}; \varphi_2^{'(2)}, \ldots \varphi_n^{'(3)}), \qquad \sigma = 2, 3 \\ \dot{A}_\sigma^{(j)} =& - \varepsilon \frac{\partial h_1}{\partial \varphi_\sigma^{'(j)}} (A_1^{'(1)}, \ldots, A_3^{'(n)}; \varphi_2^{'(2)}, \ldots \varphi_n^{'(3)}), \qquad \sigma = 2, 3 \\ \dot{\varphi}^{'(j)} =& \frac{\partial h_0}{\partial A_1^{'(j)}} (A_1^{'(1)}, \ldots, A_1^{'(n)}) \\ &+ \varepsilon \frac{\partial h_1}{\partial A_1^{'(j)}} (A_1^{'(1)}, \ldots, A_3^{'(n)}; \varphi_2^{'(2)}, \ldots \varphi_n^{'(3)}), \end{aligned} \tag{5.10.50}$$

up to $O(\varepsilon^2)$.

Since h_1 is $\varphi_1^{'(j)}$ independent, the equations in curly brackets form a system of Hamiltonian equations parameterized by the initial data of $A_1^{'(j)}$ and with $2n$ degrees of freedom. Once they are "solved", the last of Eqs. (5.10.50) is an ordinary differential equation expressing $\dot{\varphi}_1^{'(j)}$ in terms of a known function of t and, therefore, it is "trivial".

In celestial mechanics, it sometimes happens that inside the neighborhood $W = V \times \mathcal{T}^\ell$, of interesting initial data h_1 itself can be written as $\overline{h}_1 + \mu \widetilde{h}_1$ where $\overline{h}_1$, is an integrable Hamiltonian and μ is a "small" parameter.

It will then be possible to apply again perturbation theory, Proposition 16, to study the motion of the Hamiltonian system in Eqs. (5.10.50) described by the second and third line equations, as a perturbation of a simple motion, (see problem 2 at the end of this section where a similar but simpler situation occurs).

A very interesting case when this happens is the case when the unperturbed motion of the planets that one considers is a motion in which the planets wander around orbits with small eccentricity and small inclination. The resulting parameter μ is of an order of magnitude related to the maximum eccentricity and to the maximum inclination.
(4) The representation of the planetary motion thus obtained is very suggestive: the planet keeps moving with roughly the same revolution period on the same elliptic orbit [in Eqs. (5.10.50), the first and the last equations say that the average anomalies rotate with about the same unperturbed pulsations up to $O(\varepsilon)$; but the node lines and the major semiaxis longitude have a movement developing on a very slow time scale of $O(\varepsilon^{-1})$ because of the factor ε in the curly bracket equations in Eqs. (5.10.50)] called "precession" which is quasi-periodic with the periods characteristic of the Hamiltonian $\overline{h}_1$. [25] In the same quasi-periodic way vary the inclinations of the orbits and the areal velocities. The main motion obtained by neglecting $O(\varepsilon)$ in Eqs. (5.10.50) should be called a "deferent motion", while the $O(\varepsilon)$ corrections expressed by the $\sigma = 2, 3$ differential equations in Eqs. (5.10.50) should be called the "epicyclical" motions, to do some justice to the Greek astronomers and to Ptolemy, in particular.

The above "Ptolemaic" description is accurate only to $O(\varepsilon^2 + \varepsilon\mu)T$ if T is the time for which one wishes to make astronomical predictions.

The reader should consult books on celestial mechanics to see concrete applications of the procedures and approximation schemes to some astronomical problems (among which the simplest is the theoretical calculation of the precession of the perihelion of Mercury).

5.10.1 Exercises and Problems

1. Apply the idea of the proof of Proposition 17 to study the Hamiltonian system

$$\boldsymbol{\omega}_0 \cdot \mathbf{A} + \varepsilon\, g(\boldsymbol{\varphi}), \qquad (\mathbf{A}, \boldsymbol{\varphi}) \in \mathcal{R}^\ell \times \mathcal{T}^\ell$$

with $\boldsymbol{\omega}_0$ verifying Eq. (5.10.7). Deduce that the system is integrable for small ε (for an alternative solution to this problem, see Problem 1, p. 290, §4.8).

2. Apply the scheme suggested in Observation (3), p.473, to discuss to higher order the motion associated with the system in $\mathcal{R}^{\ell+1} \times \mathcal{T}^{\ell+1}$:

$$A + \varepsilon\,(\mathbf{B} \cdot \boldsymbol{\omega}_0 + \mu\, g(A, \mathbf{B}, \varphi, \boldsymbol{\psi}))$$

if $(A, \mathbf{B}, \varphi, \boldsymbol{\psi})$ are canonical action-angle variables $A \in \mathcal{R}, \mathbf{B} \in \mathcal{R}^\ell, \varphi \in \mathcal{T}^1, \boldsymbol{\psi} \in \mathcal{T}^\ell$ (note that for $\varepsilon = 0$, this system has only 1 frequency rather than $\ell + 1$). Explicitly calculate the "daily" and "secular" components of the motions to $O(\varepsilon^2 + \varepsilon\mu)$ after finding the secular Hamiltonian $\overline{h}_1$, see Observation (3), p.473, and assuming a "non resonance" condition on $\boldsymbol{\omega}_0$ like Eq. (5.10.7).

[25] It is called a "secular motion" since in some simple cases this time scale is of the order of centuries.

3. Same as Problem 2 for the system in $\mathcal{R}^2$: $\frac{1}{2}(p_1^2+q_1^2)+\frac{1}{2}(p_2^2+q_2^2)+\varepsilon\,(2q_1+q_2)^4$. (*Hint:* Find the action-angle variables $(A_1, A_2, \varphi_1, \varphi_2)$ when $\varepsilon = 0$ (just polar coordinates) for the two oscillators; then completely canonically change variables $A = \frac{1}{2}(A_1 + A_2), B = \frac{1}{2}(A_1 - A_2), \varphi = \varphi_1 - \varphi_2, \psi = \varphi_1 + \varphi_2$ and then apply the method of Observation (3), p.473.)

4. Same as Problem 2 for the system in $\mathcal{R}^2$: $\frac{p_1^2}{2}+\frac{p_2^2}{2}-\frac{1}{\sqrt{q_1^2+q_2^2}}+\varepsilon(q_1-q_2)$.

5. Consider the "restricted three-body problem" in $\mathcal{R}^2$:

$$H(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)=\frac{\mathbf{p}_1^2}{2m_1}+\frac{\mathbf{p}_2^2}{2m_2}-\frac{km_1}{|\mathbf{q}_1|}-\frac{km_2}{|\mathbf{q}_2|}-\varepsilon\frac{m_1m_2}{|\mathbf{q}_1-\mathbf{q}_2|},$$

$(\mathbf{p},\mathbf{q}) \in \mathcal{R}^2$. Using the results of Problem 15, p.305, §4.10, write (with patience) up to second order in the eccentricities of the two bodies the Hamiltonian in the action-angle variables corresponding to $\varepsilon = 0$); see Problem 11, p.303, §4.10. Show that if the eccentricities are neglected, together with quantities of order $O(\varepsilon^2)$, the secular motion [in the language of Observation (3), p.473] is described by the Hamiltonian

$$h_0+\varepsilon\overline{h}_1=-\frac{m_1^3k^2}{2L_1'^2}-\frac{m_2^3k^2}{2L_2'^2}-\int_0^{2\pi}\frac{d\alpha}{2\pi}\frac{\varepsilon m_1m_2}{\sqrt{a_1'^2+a_2'^2-2a_1'a_2'\cos\alpha}}$$

(where $L = m\sqrt{k}\,a$, $a = \{\text{major semiaxis}\}$; see Problem 11, p.303, §4.10) ("0-th order in the eccentricity").

6. Show that in the context of Problem 5, the secular Hamiltonian $\overline{h}_1$, of the Hamiltonian in Problem 5 is eccentricity independent even to first order in the eccentricity.
Does this mean that, to first order in the eccentricities, the Kepler ellipses remain fixed in space? (*Answer:* no.) Show that they move quasi-periodically "without full precession" (i.e., g_1, g_2 vary continuously with a small amplitude of oscillation, i.e., $< 2\pi$) to first order in the eccentricities.

7. Show that to second order in the eccentricities, the secular Hamiltonian of Problems 5 and 6 depends both on the L's and on the e's (i.e., on the G's) and has the form (without explicitly computing f_{ij},) $\overline{h}_1 = \overline{h}_1^{(0)}(L_1', L_2')+e_1'^2 f_{11}(L_1'.L_2', g_2'-g_1')+2e_1'e_2'f_{12}(L_1'.L_2', g_2'-g_1')+e_2'^2 f_{22}(L_1'.L_2', g_2'-g_1')$. Show that the above secular Hamiltonian is integrable and that it says that, if h are nontrivial, the relative position of the perihelions precesses to $O(e^2)$. (*Hint:* Use Problem 15, p.305, §4.10. Canonically change variables as

$$\gamma = g_1 + 2_2, \quad G=\frac{G_1+G_2}{2}, \quad \widetilde{\gamma}=g_2-g_1, \quad \widetilde{G}=\frac{G_1-G_2}{2}$$

and note that the Hamiltonian "effectively" takes the form of a Hamiltonian for a one-dimensional system (integrable by quadratures or by the Hamilton-Jacobi method).)

8. In the context of Problem 7, attempt a concrete computation of f_{ij} and of the angular velocity of the precession, assuming that the unperturbed motions take place on ellipses of small eccentricity and with semiaxes a_1, a_2 such that $a_2 - a_1$ is "of the order" of a_1 and a_2 (i.e., with quite different semiaxes).

9. (i) Let $\Gamma(L) \subset \mathcal{R}^\ell$ be a cube centered at the origin and with side $2L$. Let $\boldsymbol{\nu} \in \mathcal{Z}^\ell$, $|\boldsymbol{\nu}| > 0$ and let $\Gamma_\varepsilon(L)$ be the set of the points $\boldsymbol{\omega} \in \Gamma(L)$ such that $\frac{|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|}{|\boldsymbol{\nu}|} < \varepsilon$. Show that the measure of $\Gamma_\varepsilon(L)$ does not exceed $\varepsilon\sqrt{\ell}(2L\sqrt{\ell})^{\ell-1}$. (*Hint:* Just look at the geometrical meaning of the inequality $\frac{|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|}{|\boldsymbol{\nu}|} < \varepsilon$, the $\sqrt{\ell}$ arises from $|\boldsymbol{\nu}| = \sum_{i=1}^\ell |\nu_i| \le \sqrt{\ell}(\sum_{i=1}^\ell |\nu_i|^2)^{\frac{1}{2}}$.)

(ii) Deduce that the measure of the set Γ_C of the points $\boldsymbol{\omega} \in \Gamma(L)$ such that $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|^{-1} \le C|\boldsymbol{\nu}|^{\ell}$, $\forall\,\boldsymbol{\nu} \ne \mathbf{0}$, has a complement with Lebesgue measure not exceeding

$$2C^{-1}(2L\sqrt{\ell})^{\ell-1}\sqrt{\ell}\sum_{|\boldsymbol{\nu}|>0}\frac{1}{|\boldsymbol{\nu}|^{\ell-1}}$$

(see, also, Problem 11).

10. Using Problem 9 show that $\cup_C \Gamma_C = \widetilde{\Gamma} \subset \Gamma(L)$ has the same Lebesgue measure of $\Gamma(L)$, i.e., $(2L)^{\ell}$, although its complement is dense.

11. Without using the Lebesgue-measure theory, infer from the inequalities of Problem 9 above that $\widetilde{\Gamma}$, in Problem 10, is a dense set in $\Gamma(L)$.

12. Consider a time-dependent Hamiltonian with one degree of freedom: $h_0(A)+\varepsilon f_0(A,\varphi,t)$, where $(A,\varphi) \in \mathcal{R}^1 \times \mathcal{T}^1$ and $t \in\in \mathcal{T}^1$ is interpreted as the time appearing in a 2π-periodic time-dependent perturbation to the system with Hamiltonian h_0.
Develop a formal perturbation theory for the above system proving propositions analogous to Propositions 16 and 17 of this section. (*Hint:* Use a time-dependent canonical transformation with generating function $A'\varphi + \Phi_0(A',\varphi,t)$ and proceed, as in this section, using the Hamilton-Jacobi method.)

13. Consider the time-dependent system on $\mathcal{R}^1 \times \mathcal{T}^1$, $\frac{A^2}{2} + \varepsilon\,(\cos\varphi + \cos(\varphi - t))$, and applying the results of Problem 12, remove the perturbation to $O(\varepsilon^2)$ near the points with $A = \omega_0 = \frac{1}{2}(1+\sqrt{5})$ (see exercises and problems to §2.20 for the theory of the number ω_0).

14. Same as in Problem 13, but to $O(\varepsilon^4)$. (*Warning:* The calculations are quite long.)

15. Let $h(\mathbf{A})$ be a C^{∞} function defined on a sphere $S_{\varrho}(\mathbf{A}_0) \subset \mathcal{R}^{\ell}$ with gradient $\boldsymbol{\omega}(\mathbf{A}) = \frac{\partial h(\mathbf{A})}{\partial \mathbf{A}}$ bounded by $|\boldsymbol{\omega}(\mathbf{A})| < E$ and such that the matrix $M_{ij} = \frac{\partial^2 h}{\partial A_i \partial A_j}$ is invertible for all $\mathbf{A} \in S_{\varrho}(\mathbf{A}_0)$ and $\sum_{i,j=1}^{\ell} |(M^{-1})_{ij}| \le \eta < +\infty$. Suppose that the correspondence $\mathbf{A} \to \boldsymbol{\omega}(\mathbf{A})$ is one to one between $S_{\varrho}(\mathbf{A}_0)$ and $\boldsymbol{\omega}(S_{\varrho}(\mathbf{A}_0))$. Denote, for $C > 0$:

$$S_{\varrho}(\mathbf{A}_0, C) = \{\,\mathbf{A} \mid A \in S_{\varrho}(\mathbf{A}_0),\ |\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}|^{-1} < C|\boldsymbol{\nu}|^{\ell},\ \forall\ |\boldsymbol{\nu}| > 0\}$$

Show that there is $B > 0$, depending only on ℓ, such that

$$1 \ge \frac{\operatorname{vol} S_{\varrho}(\mathbf{A}_0, C)}{\operatorname{vol} S_{\varrho}(\mathbf{A}_0)} \ge 1 - \frac{B\,(E\eta\varrho^{-1})^{\ell}}{EC}$$

(*Hint:* Use the change of variable formula:

$$\begin{aligned}
&\int_{S_{\varrho}(\mathbf{A}_0,C)} d\mathbf{A} \equiv \int_{S_{\varrho}(\mathbf{A}_0)} d\mathbf{A} - \int_{S_{\varrho}(\mathbf{A}_0)/S_{\varrho}(\mathbf{A}_0,C)} d\mathbf{A}\\
&= \operatorname{vol}(S_{\varrho}(\mathbf{A}_0)) - \int_{\boldsymbol{\omega}(S_{\varrho}(\mathbf{A}_0))/S_{\varrho}(\mathbf{A}_0,C)} |\det\frac{\partial \mathbf{A}}{\partial \boldsymbol{\omega}}|\,\delta\boldsymbol{\omega}\\
&\ge \operatorname{vol}(S_{\varrho}(\mathbf{A}_0)) - \eta^{\ell}\int_{\boldsymbol{\omega}(S_{\varrho}(\mathbf{A}_0)/S_{\varrho}(\mathbf{A}_0,C))} d\boldsymbol{\omega}\\
&\ge \operatorname{vol}(S_{\varrho}(\mathbf{A}_0)) - \varepsilon^{\ell}\sum_{\boldsymbol{\nu}\ne\mathbf{0}}\int_{\substack{|\boldsymbol{\omega}|<E\\ |\boldsymbol{\omega}\cdot\boldsymbol{\nu}|/|\boldsymbol{\nu}|<C^{-1}|\boldsymbol{\nu}|-\ell-1}} d\boldsymbol{\omega}\\
&\ge \operatorname{vol}(S_{\varrho}(\mathbf{A}_0)) - \eta^{\ell}C^{-1}(2E\sqrt{\ell})^{\ell-1}\sqrt{\ell}\sum_{|\boldsymbol{\nu}|>0}\frac{1}{|\boldsymbol{\nu}|^{\ell+1}}
\end{aligned}$$

and then recall that $vol S_{\varrho}(\mathbf{A}_0) = const\varrho^{\ell}$.)

16. Let $\boldsymbol{\omega} = (\omega, 1) \in \mathcal{R}^2$ be such that $|\omega\nu_1 + \nu_2| \le C\,(|\nu_1| + |\nu_2|)^\alpha$ for some $\alpha, C > 0$. Let f be a function on $\mathcal{T}^1$ with Fourier coefficients $f_\nu \neq 0, \forall\,\nu \neq 0$, e.g., $f(\varphi) = 2\sum_{n=1}^{\infty} e^{-\xi n}\cos n\varphi$, $\quad \xi > 0$. Consider the Hamiltonian system on $\mathcal{R}^2 \times \mathcal{T}^2$: $H_\varepsilon = (\omega_1 A_1 + A_2) + (\varepsilon_2 + f(\varphi_1)f(\varphi_2))$. Show that the Birkhoff formal series (5.10.42) are

$$h_\varepsilon(\mathbf{A}') = (\omega A_1 + A_2) + \varepsilon\,(A_2 + f(\varphi_1)f(\varphi_2)), \qquad \text{and}$$
$$\Phi_\varepsilon(\mathbf{A}', \boldsymbol{\varphi}) = \sum_{k=1}^{\infty} \varepsilon^k \Big(\sum_{\boldsymbol{\nu}\in\mathcal{Z}^2, \boldsymbol{\nu}\neq\mathbf{0}} \frac{e^{-\xi|\boldsymbol{\nu}|} e^{i\,\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i\,(\omega\nu_1 + \nu_2)} \Big(\frac{-\nu_2}{\omega\nu_1 + \nu_2}\Big)^{k-1} \Big),$$

and prove that the series for Φ_ε does not converge. (*Hint:* Using the explicit solubility of the equations for H_ε, see Problem 1, §4.8, p.290, one sees that the passage to action-angle variables for H_ε must be singular for a dense set of values of ε: the singularities arise in correspondence of the values of ε for which the formal sum of the Φ_ε-series makes no sense (to sum formally the series permute them).)

17. In the context of Problem 16, show that the function Φ_ε, obtained by permuting $\sum_k$ and $\sum_\nu$ and summing the geometric series, makes sense and is analytic in A, φ for many values of ε and, whenever this happens, H_ε is indeed integrable by the canonical map generated by Φ_ε. (*Hint:* Use Problem 9 above to identify the values of ε which allow bounds of the type $|\omega\nu_1 + (1+\varepsilon)\nu_2|^{-1} \le C\,|\nu|^\alpha$, $C, \alpha > 0, |\nu| > 0$.)

5.11 Some Simple Properties of Holomorphic Functions. Analytic Theorems for the Implicit Functions

In §5.10, we mentioned, without discussion, some properties of the analytic functions. Such properties can be derived in the more general context of the theory of holomorphic functions.

Such functions are basically defined as analytic functions of complex variables, i.e., a $\mathcal{C}^p$-valued function f defined on an open subset $W \subset \mathcal{C}^\ell$ is holomorphic if it can be developed in an absolutely convergent power series around each point of W.

For a more detailed discussion of some perturbation theory problems, it is convenient to state the following definition which is general enough for our purposes. It is a definition that is provided more with the aim of fixing some notations rather than with the objective of developing the part of the holomorphic functions theory that we need. In this and in the following sections, we suppose that the reader is familiar with the basic properties of holomorphic functions, i.e., the Cauchy integral formula the theory of the Taylor-Laurent expansions in power series, and the identity principle. Such properties will be repeatedly used in §5.11 and §5.12.

Let ℓ, p, q be positive integers.

8 Definition. *(i) We introduce the following notations:* $\forall\,\mathbf{a} = (a_1, \ldots, \alpha_\ell) \in \mathcal{Z}_+^\ell$, $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_ell) \in \mathcal{Z}^\ell$,

$$|\mathbf{a}| = \sum_{i=1}^{\ell} |a_i|, \qquad |\boldsymbol{\nu}| = \sum_{i=1}^{p} |\nu_i|. \tag{5.11.1}$$

while if $\mathbf{w} = (w_1, \ldots, w_\ell) \in \mathcal{C}^\ell$

$$|\mathbf{w}| = \max_{1\le i\le q} |w_i|, \quad ||\mathbf{w}|| = \sum_{i=1}^{q} |w_i|. \tag{5.11.2}$$

(ii) For $\mathbf{A}_0 \in \mathcal{C}^\ell, \varrho > 0, \xi > 0$ *and* $j = 1, \ldots, p$ *we set*

$$\begin{aligned}\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0) &= \{\, \mathbf{A} \,|\, \mathbf{A} \in \mathcal{C}^\ell,\ |\mathbf{A} - \mathbf{A}_0| < \varrho \,\} \\ C(\xi) &= \{\, \mathbf{z} \,|\, \mathbf{z} \in \mathcal{C}^p,\ e^{-\xi} < |z_j| < e^{\xi} \,\}, \\ C(\varrho, \xi; \mathbf{A}_0) &= \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0) \times C(\xi).\end{aligned} \tag{5.11.3}$$

The first two such sets will be called, respectively, the "complex multisphere" with center $\mathbf{A}_0$ *and radius* ϱ*, and the "complex multiannulus", with inner radius* $e^{-\xi}$ *and outer radius* e^{ξ}*. If* $\mathbf{A}_0$ *is real, we define*

$$\mathcal{S}_\varrho(\mathbf{A}_0) = \{\, \mathbf{A} \,|\, \mathbf{A} \in \mathcal{R}^\ell,\ |A_i - A_{0i}| < \varrho \,\} \tag{5.11.4}$$

calling it the "real multisphere" with center $\mathbf{A}_0$ *and radius* ϱ*.*
The set $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ *will be identified to a subset of* $C(\varrho, \xi; \mathbf{A}_0)$ *via the map*

$$(\mathbf{A}, \boldsymbol{\varphi}) \rightarrow (\mathbf{A}, \mathbf{z}), \qquad z_j = e^{i\,\varphi_j} \tag{5.11.5}$$

(iii) If $W \subset \mathcal{C}^q$ *is open and if* F *is a* $\mathcal{C}^p$*-valued function, we say that* F *has a convergent power series expansion around* $\mathbf{w}_0 \in W$ *if there is a family of* $\mathcal{C}^p$*-vectors* $\{F^{(\mathbf{a})}(\mathbf{w}_0)\}_{\mathbf{a}\in\mathcal{Z}^q_+}$ *such that for some* $\widetilde{\varrho} > 0$*:*

$$F(\mathbf{w}) = \sum_{\mathbf{a}\in\mathcal{Z}^q_+} F^{(\mathbf{a})}(\mathbf{w}_0)(\mathbf{w} - \mathbf{w}_0)^{\mathbf{a}}, \qquad \forall\ |\mathbf{w} - \mathbf{w}_0| < \widetilde{\varrho}, \tag{5.11.6}$$

having set

$$(\mathbf{w} - \mathbf{w}_0)^{\mathbf{a}} = \prod_{j=1}^{q} (w_j - w_{0j})^{a_j}, \qquad \textit{for } \mathbf{a} = (a_1, \ldots, a_q), \textit{ and} \tag{5.11.7}$$

$$\sum_{\mathbf{a}\in\mathcal{Z}^q_+} |F^{(\mathbf{a})}(\mathbf{w}_0)|\, \varrho^{|\mathbf{a}|}(\mathbf{w}_0), \qquad \forall\,\forall\, \varrho < \widetilde{\varrho} \tag{5.11.8}$$

(iv) A function F *is holomorphic in the open subset* $W \subset \mathcal{C}^q$ *if it has a convergent power series around every point* $\mathbf{w} \in W$*. In this case, one defines the derivatives of* F *as*

$$\frac{\partial^{|\mathbf{a}|} F(\mathbf{w}_0)}{\partial \mathbf{w}^{\mathbf{a}}} = \mathbf{a}!\, F^{(\mathbf{a})}(\mathbf{w}_0), \qquad \forall\ \mathbf{w}_0 \in W \tag{5.11.9}$$

if $\mathbf{a}! \stackrel{def}{=} \prod_{j=1}^{q} a_j!$.

Observation. One calls Eq. (5.11.6) the Taylor series of F at $\mathbf{w}_0$, because of Eq. (5.11.9).

9 Definition. *Let ℓ, p, q be positive integers and $\mathbf{A}_0 \in \mathcal{R}^\ell$ and use the notations of Definition 8 above.*
(i) Let f, g, h be three functions defined, respectively, on $\mathcal{S}_\varrho(\mathbf{A}_0), \mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p, \mathcal{T}^p$ with values in $\mathcal{R}^q$. We shall say that they are holomorphic in $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, $C(\varrho, \xi; \mathbf{A}_0)$, $C(\xi)$ respectively if, identifying $\mathcal{S}_\varrho(\mathbf{A}_0), \mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p, \mathcal{T}^p$ as subsets of $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0), C(\varrho, \xi; \mathbf{A}_0), C(\xi)$, as explained in Definition 8 (ii) above, they can be extended to holomorphic functions $\overline{f}, \overline{g}, \overline{h}$ on the larger sets $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0), C(\varrho, \xi; \mathbf{A}_0), C(\xi)$.
The functions of the type $\overline{f}, \overline{g}, \overline{h}$ will be called "holomorphic" in $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, $C(\varrho, \xi; \mathbf{A}_0), C(\xi)$, respectively, and "real" on $\mathcal{S}_\varrho(\mathbf{A}_0), \mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p, \mathcal{T}^p$, respectively. Sometimes the extensions $\overline{f}, \overline{g}, \overline{h}$ will still be called f, g, h, dropping the bar.
(ii) If F is holomorphic on $C(\varrho, \xi; \mathbf{A}_O)$ or on $C(\xi)$, we define its "$\boldsymbol{\varphi}$-derivatives" by setting $\frac{\partial}{\partial \varphi_k} = i z_k \frac{\partial}{\partial z_k}$, $k = 1, \ldots, p$.

Observations.
(1) It is easy to deduce from the definition of an analytic function on $V \times \mathcal{T}^\ell$ (see Definitions 13 and 14, p.336 and p.337, §4.13) that if f, g and h are analytic on V or on $V \times \mathcal{T}^p, \mathcal{T}^p$, respectively, then given $\mathbf{A}_0 \in V$, there exist $\varrho, \xi > 0$ such that f is holomorphic in $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, g in $C(\varrho, \xi; \mathbf{A}_O)$, and h in $C(\xi)$. In general, however, ϱp and ξ may be very small even if V is large.
(2) This definition is particularly useful because it provides a simple description of an important class of functions on $\mathcal{T}^p$ or on $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p$, thinking of $\mathcal{T}^p$ as a subset of $C(\xi)$ via the natural correspondence

$$\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p) \in \mathcal{T}^p \longleftrightarrow \mathbf{z} = (e^{i\varphi_1}, \ldots, e^{i\varphi_q}) \tag{5.11.10}$$

already pointed out several times.

The classical theorems on the theory of the holomorphic functions (Taylor and Laurent expansions, Cauchy's formula, identity principle, etc.) imply the following proposition which we do not prove since it can be found, with other symbols, in any elementary textbook on holomorphic functions.

18 Proposition. *Let f, g, h be holomorphic functions on $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, $C(\varrho, \xi; \mathbf{A}_0)$, $C(\xi)$, respectively, respectively, see Eq. (5.11.4), with values in $\mathcal{C}^q$. Using the notation of Definitions 8 and 9 and setting*

$$\mathbf{z} = \prod_{j=1}^{p} z^{\nu_j}, \qquad \text{for } \boldsymbol{\nu} \in \mathcal{Z}^p, \mathbf{z} \in C(\xi) \tag{5.11.11}$$

(i) Sequences of vectors in $\mathcal{C}^q$ $\{f^{(\mathbf{a})}\}_{\mathbf{a}\in\mathcal{Z}^\ell_+}$, $\{g^{(\mathbf{a})}_{\boldsymbol{\nu}}\}_{\mathbf{a}\in\mathcal{Z}^\ell_+,\boldsymbol{\nu}\in\mathcal{Z}^p}$, $\{h_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\mathcal{Z}^p}$ *exist such that*

$$\begin{aligned}\overline{f}(\mathbf{A}) &= \sum_{\mathbf{a}\in\mathcal{Z}^\ell_+} f^{(\mathbf{a})}\,(\mathbf{A}-\mathbf{A}_0)^{\mathbf{a}},\\ g(\mathbf{A},\mathbf{z}) &= \sum_{\mathbf{a}\in\mathcal{Z}^\ell_+,\boldsymbol{\nu}\in\mathcal{Z}^p} g^{(\mathbf{a})}_{\boldsymbol{\nu}}\,(\mathbf{A}-\mathbf{A}_0)^{\mathbf{a}}\mathbf{z}^{\boldsymbol{\nu}},\\ h(\mathbf{z}) &= \sum_{\boldsymbol{\nu}\in\mathcal{Z}^p} h_{\boldsymbol{\nu}}\,\mathbf{z}^{\boldsymbol{\nu}}.\end{aligned}\tag{5.11.12}$$

(ii) Identifying $g(\mathbf{A},\boldsymbol{\varphi})$ *as* $g(\mathbf{A},\mathbf{z})$, *and* $h(\boldsymbol{\varphi})$ *with* $h(\mathbf{z})$ *for all* $\mathbf{z}=(e^{i\varphi_1},\ldots,e^{i\varphi_p})$, $\boldsymbol{\varphi}\in\mathcal{T}^p$, *then*

$$\begin{aligned}f^{(\mathbf{a})} &= \frac{1}{\mathbf{a}!}\frac{\partial^{|\mathbf{a}|}f}{\partial\mathbf{A}^{\mathbf{a}}}(\mathbf{A}_0),\\ g^{(\mathbf{a})}_{\boldsymbol{\nu}} &= \frac{1}{\mathbf{a}!}\int_{\mathcal{T}^p}\frac{\partial^{|\mathbf{a}|}g}{\partial\mathbf{A}^{\mathbf{a}}}(\mathbf{A}_0,\boldsymbol{\varphi})e^{-i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}\frac{d\boldsymbol{\varphi}}{(2\pi)^p},\\ h_{\boldsymbol{\nu}} &= \int_{\mathcal{T}^p} h(\boldsymbol{\varphi})e^{-i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}\frac{d\boldsymbol{\varphi}}{(2\pi)^p}.\end{aligned}\tag{5.11.13}$$

(iii) Setting

$$|f|_\varrho = \sup|f(\mathbf{A})|,\ |g|_{\varrho,\xi} = \sup|g(\mathbf{A},\mathbf{z})|,\ |h|_\xi = \sup|h(\mathbf{z})|,\tag{5.11.14}$$

where the suprema are taken over the functions respective domains of definition [and Eq. (5.11.2) is used for $||\cdot||$*], it is*

$$|f^{(\mathbf{a})}|\le|f|_\varrho\varrho^{-|\mathbf{a}|},\ |g^{(\mathbf{a})}_{\boldsymbol{\nu}}|\le|g|_{\varrho,\xi}\varrho^{-|\mathbf{a}|}e^{-\xi|\boldsymbol{\nu}|},\ |h_{\boldsymbol{\nu}}|\le e^{-\xi|\boldsymbol{\nu}|}.\tag{5.11.15}$$

(iv) If the coefficients of the series in Eq. (5.11.12) can be bounded by a constant times, respectively, $\varrho^{-|\mathbf{a}|}$, *or* $\varrho^{-|\mathbf{a}|}$, *or* $e^{-\xi|\boldsymbol{\nu}|}$, *then the sums of the series of Eq. (5.11.12) define holomorphic functions on* $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, $C(\varrho,\xi;\mathbf{A}_0)$, $C(\xi)$, *respectively.*
(v) The second of Eqs. (5.11.12) can also be written

$$g(\mathbf{A},\mathbf{z}) = \sum_{\boldsymbol{\nu}\in\mathcal{Z}^p} g_{\boldsymbol{\nu}}(\mathbf{A})\,\mathbf{z}^{\boldsymbol{\nu}}\tag{5.11.16}$$

with $g_{\boldsymbol{\nu}}(\mathbf{A})$ *holomorphic in*$\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$ *and such that its Taylor series around* $\mathbf{A}_0$ *is obtained by inspecting the second of Eqs. (5.11.12) and considering the sum over* $\mathbf{a}$ *only. (vi) If* f,g,h *are real on* $\mathcal{S}_\varrho(\mathbf{A}_0)$, $\mathcal{S}_\varrho(\mathbf{A}_0)\times\mathcal{T}^p$, $\mathcal{T}^p$ *for* $\mathbf{A}_0\in\mathcal{R}^\ell$, *the* $f^{(\mathbf{a})}$ *coefficients are also real, while* $g^{(\mathbf{a})}_{\boldsymbol{\nu}}$ *and* $h_{\boldsymbol{\nu}}$ *complex conjugates to* $g^{(\mathbf{a})}_{-\boldsymbol{\nu}}$ *and* $h_{-\boldsymbol{\nu}}$ *respectively, and vice versa.*

Observations.
(1) Note that the convergence of the series of Eq. (5.11.12) stated in (i) follows from Eq. (5.11.15) only if $|f|_\varrho$, $|g|_{\varrho,\xi}$, and $|h|_\xi$ are finite. This is, however, not necessarily true in general so that (iii) and (iv) are not reciprocal statements.
(2) If F is holomorphic in a region $W \subset \mathcal{C}^q$ and $\widetilde{\mathbf{w}} \in W, \widehat{\mathcal{S}}_\varrho(\widetilde{\mathbf{w}}) \subset W$, and if we wish to estimate the derivatives $\frac{\partial F(\widetilde{\mathbf{w}})}{\partial w_k}$, we can use Eqs. (5.11.15) and (5.11.13) as follows.
Here and below, we regard a matrix-valued function with values on the matrices $\ell \times q$ as a $\mathcal{C}^{\ell q}$-valued function,[26] and consider F as a holomorphic function on $\widehat{\mathcal{S}}_\varrho(\widetilde{\mathbf{w}})$. To bound the $\ell \times q$ matrix $\frac{\partial F}{\partial \mathbf{w}}$ (assuming that F is C^ℓ-valued), consider the first of Eqs. (5.11.13) and (5.11.15) with $|\mathbf{a}| = 1$. It gives

$$|\frac{\partial F}{\partial \mathbf{w}}(\widetilde{\mathbf{w}})| \le \big(\sup_{\mathbf{w}\in\widehat{\mathcal{S}}_\varrho(\widetilde{\mathbf{w}})} |F(\mathbf{w})|\big)\varrho^{-1} \le \frac{\sup|F(\mathbf{w})|}{\varrho}, \tag{5.11.17}$$

where the second supremum is over W. From this remark, it follows that

$$\begin{aligned} |\frac{\partial f}{\partial \mathbf{A}}|_{\varrho'} \le & \frac{|f|_\varrho}{\varrho-\varrho'}, \quad |\frac{\partial g}{\partial \mathbf{A}}|_{\varrho',\xi} \le \frac{|g|_{\varrho,\xi}}{\varrho-\varrho'}, \\ |\frac{\partial g}{\partial \mathbf{z}}|_{\varrho,\xi'} \le & \frac{|g|_{\varrho,\xi}}{e^{-\xi'}-e^{-\xi}} \le |g|_{\varrho,\xi}\frac{e^\xi}{\delta}, \\ |\frac{\partial g}{\partial \varphi_k}|_{\varrho,\xi'} \equiv & |iz_k \frac{\partial g}{\partial z_k}|_{\varrho,\xi'} \le |g|_{\varrho,\xi}\frac{e^{2\xi}}{\delta} \end{aligned} \tag{5.11.18}$$

for $\varrho' < \varrho, \xi' < \xi$, if $\delta = \xi - \xi'$. Analogous inequalities hold for the higher order derivatives, e.g., $|\frac{\partial^2 f}{\partial \mathbf{A}\partial \mathbf{A}}|_{\varrho'} \le 2\frac{|f|_\varrho}{(\varrho-\varrho')^2}$ [see Eq. (5.11.9)].

These simple estimates will be called "dimensional estimates". In physics, one says that a "dimensional estimate" is any estimate of the derivative of a function F at a given point in terms of the function maximum in a region divided by the distance of the point to the region boundary ("characteristic magnitude of F" divided by a "characteristic length"). Recall that physicists rightly believe that all functions (with, possibly, some exceptions) are analytic.

We now possess the terminology necessary to formulate the analytic implicit function theorem. This theorem is a particularly simple and strong version of the ordinary implicit function theorem valid when the defining function is analytic. It will play a key role in the proof of Proposition 22 which, in turn, is the heart of the proof of Proposition 15, p.460, (the KAM theorem) in the analytic case.

The proof of the propositions that follow uses elementary aspects of the theory of holomorphic functions and it will be discussed in Appendix N. Propositions 19-21 are "analytic implicit function theorems".

---
[26] so that $|M(\mathbf{w})| = \sup_{ij}|M_{ij}(\mathbf{w})|$, $||M(\mathbf{w})|| = \sum_{ij}|M_{ij}|$.

19 Proposition. *Let $\ell > 0$ be an integer, $\mathbf{A}_0 \in \mathcal{R}^\ell$, and f be a $\mathcal{C}^\ell$-valued function holomorphic in the complex multisphere $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$ and real on $\mathcal{S}_\varrho(\mathbf{A}_0)$. Consider the equation for $\mathbf{A}$:*

$$\mathbf{A} - \mathbf{A}_0 + \mathbf{f}(\mathbf{A}) = \mathbf{0}. \tag{5.11.19}$$

There exists a constant γ (one can take, e.g., $\gamma = 2^8$) such that if

$$\gamma\,|\mathbf{f}|_\varrho < 1, \tag{5.11.20}$$

Eq. (5.11.19) admits a unique solution $\mathbf{A}_1 \in \mathcal{S}_\varrho(\mathbf{A}_0)$, i.e. $\mathbf{A}_1 \in \mathcal{R}^\ell$ and

$$|\mathbf{A}_1 - \mathbf{A}_0| < \varrho. \tag{5.11.21}$$

A corresponding proposition can be formulated for equations in $\mathcal{T}^\ell$.

20 Proposition. *Let $p, \ell > 0$ be integers, let $\mathbf{A}_0 \in \mathcal{R}^\ell$, and $\varrho, \xi, \delta > 0$. Let $\mathbf{g}$ be an $\mathcal{R}^p$-valued analytic function on $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p$ holomorphic in $C(\varrho, \xi; \mathbf{A}_O)$. Consider the equation*

$$\boldsymbol{\varphi}' = \boldsymbol{\varphi} + \mathbf{g}(\mathbf{A}, \boldsymbol{\varphi}) \tag{5.11.22}$$

thought of as an equation on $\mathcal{T}^p$ parameterized by $\boldsymbol{\varphi}' \in \mathcal{T}^p$ and $\mathbf{A} \in \mathcal{S}_\varrho(\mathbf{A}_0)$. Then there exists a constant γ (e.g., $\gamma = 2^8$) such that:
(i) Equation (5.11.22) is soluble if

$$\gamma\,|\mathbf{g}|_{\varrho,\xi}\, e^{2\xi} \delta^{-1} < 1 \tag{5.11.23}$$

and admits a solution of the form

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}, \boldsymbol{\varphi}') \tag{5.11.24}$$

with $\boldsymbol{\Delta}$ being an $\mathcal{R}^p$-valued analytic function on $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p$ holomorphic in $C(\varrho, \xi, \mathbf{A}_0)$.
(ii) The function $\boldsymbol{\Delta}$ can be bounded as

$$|\boldsymbol{\Delta}|_{\varrho,\xi-\delta}| \le |\mathbf{g}|_{\varrho,\xi}. \tag{5.11.25}$$

(iii) The only function inverting Eq. (5.11.22) and enjoying the properties (i) and (ii) above is $\boldsymbol{\Delta}$.

Observations.
(1) The reader should note that the above two implicit function theorems have “dimensional nature”, i.e., they just say what can be naively guessed.

In fact, in order to invert an implicit equation “close to the identity” like Eq. (5.11.22), one expects to have to impose that the derivatives of $\mathbf{g}$ are small compared to the derivatives of the identity map (i.e., small compared to 1). This is precisely the meaning of Eq. (5.11.23): if we wish to invert inside the annulus with external radius $e^{\xi+\delta}$ and internal radius $e^{\xi-\delta}$, we estimate the

gradient of $\mathbf{g}$ in the region by $|\mathbf{g}|_{\varrho,\xi}\delta^{-1}e^{\xi}$, see Eq. (5.11.18). For $\xi \gg 1$, this is still not the same as Eq. (5.11.23) (while it is such for $\xi \le 1$). However, if ξ is large, we are asking for the inversion of Eq. (5.11.16) in a very large region and extra conditions stem out of the requirement of global invertibility[27] (see the proof).
(2) Proposition 19 is an infinite-dimensional version of the implicit function theorem, since one can consider all the Taylor coefficients of $\mathbf{f}$ at $\mathbf{A}_0$ as parameters in Eq. (5.11.19). Also, Eq. (5.11.22) is susceptible to such an interpretation.
(3) Note that the constant γ in Propositions 19 and 20 is ℓ and p-independent. It is also the same in Propositions 19 and 20: but (this has been arranged so as to avoid introducing too many (constants). The numerical value of γ is not optimal.
(4) Proposition 20 is remarkable because it is a "global inversion" theorem. The equation is posed on all of $\mathcal{T}^p$ and not just locally.

A proposition analogous to Proposition 20 holds for the equation

$$\mathbf{w}' = \mathbf{w} + \mathbf{G}(\mathbf{w}, \mathbf{z}), \tag{5.11.26}$$

where $\mathbf{G}$ is a $\mathcal{C}^\ell$-valued function holomorphic on $C(\varrho, \xi; \mathbf{A}_0)$ and real on $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p \subset C(\varrho, \xi; \mathbf{A}_0)$:

21 Proposition. *Let $\ell, p > 0$ be integers, $\mathbf{A}_0 \in \mathcal{R}^\ell$ and $\varrho, \xi, \tau > 0$, $\tau < 1$. Let $\mathbf{G}$ be an $\mathcal{R}^\ell$-valued analytic function on $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p$ holomorphic on $C(\varrho, \xi; \mathbf{A}_0)$.*
Consider Eq. (5.11.26) as an equation for $\mathbf{w}$ parameterized by $\mathbf{w}', \mathbf{z}$.
(i) There is a constant γ (e.g., again, $\gamma = 2^8$) such that if

$$\gamma\,|\mathbf{G}|_{\varrho,\xi}\varrho^{-1}\tau^{-1} < 1, \tag{5.11.27}$$

Eq. (5.11.26) is soluble, $\forall\, \mathbf{w}' \in \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$ and admits a solution of the form

$$\mathbf{w} = \mathbf{w}' + \mathbf{D}(\mathbf{w}', \mathbf{z}) \tag{5.11.28}$$

with $\mathbf{D}$ holomorphic on $C(\varrho e^{-\tau}, \xi; \mathbf{A}_0)$ real on $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^p$.
(ii) The following bound can be put on $\mathbf{D}$:

$$|\mathbf{D}|_{\varrho e^{-\tau},\xi} \le |\mathbf{G}|_{\varrho,\xi}. \tag{5.11.29}$$

(iii) $\mathbf{D}$ is the only function inverting Eq. (5.11.26) and enjoying the properties (i) and (ii) above.
(iv) Fixing $\mathbf{w} \in C(\varrho e^{-\tau}, \xi; \mathbf{A}_0)$, Eq. (5.11.28) yields the only $\mathbf{w} \in C(\varrho, \xi; \mathbf{A}_0)$ verifying Eq. (5.11.26).

Observations.
(1) The above proposition makes sense, and is true, in a natural way if $p = 0$

[27] Also, it makes a difference to bound $\partial/\partial\boldsymbol{\varphi}$ rather than $\partial/\partial\mathbf{z}$ for ξ large.

(just drop everywhere $\mathbf{z}$ and the index ξ). Likewise, Proposition 20 makes sense in a natural way if $\ell = 0$ (just drop $\mathbf{A}$ and the index ϱ everywhere).
(2) Setting $\mathbf{w}' = \mathbf{0}$ in Eq. (5.11.26) as well as $p = 0, \mathbf{A}_0 = \mathbf{0}$ and applying Proposition 21, one deduces Proposition 19 with $\mathbf{A}_0 = \mathbf{0}$. Since $\mathbf{A}_0 = \mathbf{0}$ is clearly not restrictive, Proposition 19 is a corollary of Proposition 21.
(3) This proposition is clearly analogous to Proposition 20 and has, also, a "dimensional nature", see observation (1), p.484; more generally, the comments made on Proposition 20 can be repeated with obvious modifications for Proposition 21. An analogue of item (iv) in Proposition 21 could also be formulated for Proposition 20, but it will not be needed.

5.11.1 Problems and Exercises

1. After studying the proof in Appendix N of Proposition 20, find a better value for the constant γ.

2. Same as Problem 1 for Proposition 21 and for its corollary, Proposition 19.

3. Apply Proposition 20 to invert the equation $\ell = \xi - \varepsilon \sin\xi$, $\xi \in \mathcal{T}^1, \ell \in \mathcal{T}^1$, appearing in the theory of the two-body problem, see Problem 13, p.304, §4.10. Here ε is a parameter, $0 < \varepsilon < 1$ ("eccentricity"). Find for which values of $\varepsilon \in \mathcal{R}$ the above equation can be globally inverted if the estimates in the theorem are applied.

4. Same as Problem 3 with the new γ computed in Problem 1.

5. Show that the equation in Problem 3 can be inverted for all $\varepsilon \in [0, 1)$ in the sense that there is a function g analytic on $\mathcal{T}^1$ such that $\xi = \ell - g(\ell)$, for each given $\varepsilon \in [0, 1)$. (*Hint:* Do not use Proposition 20 directly.)

6.* (Levi-Civita) Check that $g(\ell)$ is holomorphic in the unit disk $|\eta| < 1$ if

$$\eta \stackrel{def}{=} \frac{\varepsilon\, e^{\sqrt{1-\varepsilon^2}}}{1 + \sqrt{1-\varepsilon^2}}$$

(from Vol. 2, p. 321 in [24]). Draw with the help of a computer the curve in the complex plane of the ε's such that $|\eta| = 1$ and check that its point closest to the origin is imaginary and at a distance $\varepsilon_L \sim 0.662...$ It is the radius of convergence in ε of $g(\ell)$. ("Laplace limit", p.304). (*Hint:* Given $\varepsilon \in C$ the Jacobian of the map is $1 - \varepsilon \cos\xi$ and is 0 for $\cos\xi = \frac{1}{\varepsilon}$. This means that the equation $\zeta = z e^{-\varepsilon \sin\xi}$, with $\zeta = e^{i\ell}$, $z = e^{i\xi}$, has a solution with ℓ real if $|\cos\xi + i \sin\xi| e^{-i\varepsilon \sin\xi} > 1$. Since $\cos\xi = \frac{1}{\varepsilon}$ $\sin\xi = \frac{\pm 1}{\varepsilon}\sqrt{\varepsilon^2 - 1}$ this is implied by $\eta = \frac{\varepsilon\, e^{\sqrt{1-\varepsilon^2}}}{1+\sqrt{1-\varepsilon^2}} < 1$. Check that the singularity of g in ε closest to the region occurs for $\varepsilon = i\varrho$ and $\frac{\varrho}{1+\sqrt{1+\varrho^2}} e^{\sqrt{1+\varrho^2}} = 1$, which defines the radius of convergence ε_L of g in powers of ε, i.e. the Laplace limit.)

5.12 Perturbations of Trajectories. Small Denominators Theorem

Another perturbative problem that could be studied is the following. Let $(\mathbf{A}, \boldsymbol{\varphi}) \to h_0(\mathbf{A})$ be an analytic Hamiltonian on $V \times \mathcal{T}^\ell$ which we suppose such that the matrix

$$M_0(\mathbf{A})_{ij} = \frac{\partial^2 h_0}{\partial A_i \partial A_j}(\mathbf{A}) \tag{5.12.1}$$

has determinant $\neq 0$ on $V \times \mathcal{T}^\ell$ ("integrable non isochronous system").

Given $\mathbf{A}_0 \in V$, the torus $\{\mathbf{A}_0\} \times \mathcal{T}^\ell$ is an ℓ-dimensional torus invariant for the motion associated with the Hamiltonian h_0. The Hamiltonian flow on the phase space $V \times \mathcal{T}^\ell$ induces on the torus a quasi-periodic flow $\boldsymbol{\varphi} \to \boldsymbol{\varphi} + \boldsymbol{\omega}_0 t$, $t \geq 0$, with pulsations

$$\boldsymbol{\omega}_0 = \frac{\partial h_0}{\partial \mathbf{A}}(\ A_0) \stackrel{def}{=} \boldsymbol{\omega}(\mathbf{A}_0). \tag{5.12.2}$$

If f_0 is an analytic function on $V \times \mathcal{T}^\ell$, it is natural to ask whether the motions on $V \times \mathcal{T}^\ell$ associated with the perturbed Hamiltonian,

$$H_0(\mathbf{A}, \boldsymbol{\varphi}) = h_0(\mathbf{A}) + f_0(\mathbf{A}, \boldsymbol{\varphi}), \tag{5.12.3}$$

leave a torus invariant, inducing on it a quasi-periodic flow with pulsations $\boldsymbol{\omega}(\mathbf{A}_0)$, i.e., "with the same spectrum" as before. One could call this problem "the spectrum-conservation problem".

Intuitively, one could expect that a torus on which a quasi-periodic motion with pulsations $\boldsymbol{\omega}(\mathbf{A}_0)$ takes place will continue to exist but it will be "deformed" inside $V \times \mathcal{T}^\ell$ if compared to the one relative to the $f_0 = 0$ case, at least if $M_0(\mathbf{A}_0)$ is invertible[28] and f_0 is small.

This perturbation problem differs from the one of the preceding sections; the latter was in fact concerned with the study of the perturbations of motions with given initial datum. A whole family of motions is now considered, which enjoy a certain common property, namely, quasi-periodicity with pulsations $\boldsymbol{\omega}(\mathbf{A}_0)$, and we ask whether a family of motions with the same property still exists after perturbation. Proposition 15 of §5.9 provides an answer, in some sense affirmative.

A proposition will now be formulated which, as it appears from the observations that follow it, also proves important parts of Proposition 15 and gives all the ingredients necessary for its full proof in the analytic case.

The proof of Proposition 22 that follows is taken from Arnold and is specifically fit for the analytic case under examination. The analogous proposition

[28] In the case $\boldsymbol{\omega}(\mathbf{A}) \equiv \boldsymbol{\omega}_0$, $\forall\, \mathbf{A} \in V$ and, hence, $M_0(\mathbf{A}) \equiv 0$ it is easy to give a counterexample. Let $\ell = 1, h_0(\mathbf{A}) = A$ so that $\omega(A) \equiv 1$ and $M_0 \equiv 0$. Let $f(A, \varphi) \equiv \varepsilon A$. Then the unperturbed motions have pulsation 1, while the perturbed ones have pulsation $(1 + \varepsilon) \neq 1$, if $\varepsilon \neq 0$.

in the $C^{(k)}$-differentiable case (with k large enough) is due to Moser, [33], and is based on a technically different method.

Before stating Proposition 22, which will be called "small denominators theorem" for reasons manifest from its proof (or "Arnold's theorem", [2]), some notations are needed, see also Eqs. (5.11.1)-(5.11.3), (5.11.14).

10 Definition. *(i) If* $\mathbf{a} \in Z_+^\ell$, $\boldsymbol{\nu} \in \mathcal{Z}^\ell$, *let* $|\mathbf{a}| = \sum_{i=1}^{\ell} |a_i|$, $|\boldsymbol{\nu}| = \sum_{i=1}^{p} |\nu_i|$.
(ii) If $\mathbf{w} \in \mathcal{C}^q$, *we set* $|\mathbf{w}| = \max_{1 \le i \le q} |w_i|$, $\quad ||\mathbf{w}|| = \sum_{i=1}^{q} |w_i|$.
A $\ell \times \ell$ *matrix* M *will be regarded as an element of* $\mathcal{C}^q$ *with* $q = \ell^2$ *so that it will make sense to write* $|M|$, $||M||$.
(iii) If f, h *are holomorphic in* $C(\varrho, \xi; \mathbf{A}_0), \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$ *respectively*[29] *and take values in* $\mathcal{C}^q$ *let [see Eqs. (5.11.11) and (5.11.3)]*

$$\begin{aligned} |f|_{\varrho,\xi} &= \sup(\mathbf{A}, \mathbf{z})|, \qquad ||f(\mathbf{A}, \mathbf{z})||, \\ |h|_\varrho &= \sup |h(\mathbf{A})|, \qquad ||h||_\varrho = \sup ||h(\mathbf{A})|| \end{aligned}$$

where the suprema are taken over the domains of the various functions.

The small-denominators theorem can then be formulated as follows.

22 Proposition. *Let* h_0, f_0 *be two real analytic functions* $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ *holomorphic in* $C(\varrho_0, \xi_0; \mathbf{A}_0)$, $\xi_0 < 1$. *Assume that* h_0 *depends only on the action variables* $\mathbf{A}$ *in* $(\mathbf{A}, \boldsymbol{\varphi}) \in \mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ *and that the matrix* M_0 *of Eq. (5.12.1) is nonsingular. Suppose that* $\boldsymbol{\omega}_0 = \frac{\partial h_0}{\partial \mathbf{A}}(\mathbf{A}_0)$ *has the "non resonance" property:*

$$|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}|^{-1} \le C |\boldsymbol{\nu}|^\ell, \qquad \forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell,\ |\boldsymbol{\nu}| > 0 \tag{5.12.4}$$

for some $C > 0$ *("resonance parameter"). Let* $E_0, \eta_0, \varepsilon_0$ *be such that*

$$E_0 > |\frac{\partial h_0}{\partial \mathbf{A}}|_{\varrho_0,\xi_0},\ \eta_0 > ||M_0^{-1}||_{\varrho_0},\ \varepsilon_0 > |\frac{\partial f_0}{\partial \mathbf{A}}|_{\varrho_0,\xi_0} + \frac{1}{\varrho_0} |\frac{\partial f_0}{\partial \boldsymbol{\varphi}}|_{\varrho_0,\xi_0} \tag{5.12.5}$$

Then there exist constants $B, a, b, c > 0$, *only depending upon the number* ℓ *of degrees of freedom,*[30] *such that if*

$$q \stackrel{def}{=} B\, C\, \varepsilon_0\, (C E_0)^a\, (\eta_0 E_0 \varrho_0^{-1})^b \xi_0^{-c} < 1 \tag{5.12.6}$$

one can find in $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ *a torus* $\mathcal{T}(\boldsymbol{\omega}_0)$ *with parametric equations*

$$\begin{aligned} \mathbf{A} &= \mathbf{A}_0 + \boldsymbol{\alpha}(\boldsymbol{\varphi}'), \\ \boldsymbol{\varphi} &= \boldsymbol{\varphi}' + \boldsymbol{\beta}(\boldsymbol{\varphi}'), \end{aligned} \qquad \boldsymbol{\varphi}' \in \mathcal{T}^\ell \tag{5.12.7}$$

and such that:
(i) $\mathcal{T}(\boldsymbol{\omega}_0)$ *is invariant for the evolution in* $\mathcal{S}_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ *associated with the Hamiltonian (5.12.3). On* $\mathcal{T}(\boldsymbol{\omega}_0)$ *the evolution is described by the map*

[29] See (5.11.3) for the meaning of the symbols.
[30] e.g., a rather rough, though not "totally absurd", estimate says that one can take $a = b = 14$, $c = 2(10\ell + 6)$, $B = (12\ell)! 10^{40\ell}$ (very far from optimal).

$$\boldsymbol{\varphi}' \to \boldsymbol{\varphi}' + \boldsymbol{\omega}\, t, \qquad t \in \mathcal{R}_+ \tag{5.12.8}$$

and is therefore quasi-periodic with pulsations $\boldsymbol{\omega}_0$*.*
(ii) The functions $\boldsymbol{\alpha}, \boldsymbol{\beta}$ *are analytic on* $\mathcal{T}^\ell$ *and*

$$\varrho_0^{-1}|\boldsymbol{\alpha}(\boldsymbol{\varphi}')| + |\boldsymbol{\beta}(\boldsymbol{\varphi}')| \le q. \tag{5.12.9}$$

Observations.
(1) Using the notations of Proposition 18, p.481, §5.11, Proposition 18, Eqs. (5.11.16) and (5.11.12), imply that f_0 can be written as

$$f_0(\mathbf{A}, \mathbf{z}) = \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell} f_{0\boldsymbol{\nu}}(\mathbf{A})\, \mathbf{z}^{\boldsymbol{\nu}} = \sum_{\mathbf{a} \in \mathcal{Z}^\ell_+, \boldsymbol{\nu} \in \mathcal{Z}^\ell} f_{0\boldsymbol{\nu}}^{(\mathbf{a})}\, (\mathbf{A} - \mathbf{A}_0)^{\mathbf{a}} \mathbf{z}^{\boldsymbol{\nu}}, \tag{5.12.10}$$

where $f_{0\boldsymbol{\nu}}(\mathbf{A})$ is the sum of the series in $\mathbf{a}$ in the right-hand side of Eq. (5.12.10) and is holomorphic in $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$.
Then the derivatives $\partial/\partial\boldsymbol{\varphi}$ appearing in Eq. (5.12.5) can be simply defined as $\partial/\partial\varphi_k \stackrel{def}{=} i\, z_k \partial/\partial z_k$ (see Definition 9 (ii), p.481).

(2) It follows from the theory of the Taylor-Laurent expansions for holomorphic functions that if $g_1, \ldots, g_r$ are r real analytic functions on $V \times \mathcal{T}^\ell$, $V \subset \mathcal{R}^r$ open, it is possible to find two functions $\mathbf{A} \to \varrho(\mathbf{A})$, $\mathbf{A} \to \xi(\mathbf{A})$ positive and continuous on V such that $S_{\varrho(\mathbf{A})} \subset V$, $\forall\, \mathbf{A} \in V$, and, furthermore, such that $g_1, \ldots, g_r$ are holomorphic in $C(\varrho(\mathbf{A}), \xi(\mathbf{A}); \mathbf{A})$, see Definition 8, p.479, and $|g_j|_{\varrho(\mathbf{A}), \xi(\mathbf{A})}$ are continuous functions of $\mathbf{A}$ in V, $j = 1, \ldots, r$.
Therefore Proposition 22 could be formulated in an apparently more general form by only requiring the analyticity of h_0, f_0, M_0^{-1} in $\widehat{\mathcal{S}}_\varrho(\mathbf{A}_0) \times \mathcal{T}^\ell$ rather than their holomorphy in $C(\varrho, \xi; \mathbf{A}_0)$.

(3) An elementary result of measure theory implies that, given $C > 0$ and supposing V bounded, the set $V(C)$ of the points $\mathbf{A} \in V$ such that

$$|\boldsymbol{\nu} \cdot \boldsymbol{\omega}(\mathbf{A})|^{-1} \equiv \big|\boldsymbol{\nu} \cdot \frac{\partial h_0(\mathbf{A})}{\partial \mathbf{A}}\big|^{-1} \le C|\boldsymbol{\nu}|^\ell, \quad \forall\ \boldsymbol{\nu} \in \mathcal{Z}^\ell, |\boldsymbol{\nu}| > 0$$

has a Lebesgue measure $\mu(V(C)) \xrightarrow[C \to +\infty]{} \mu(V)$ if $M_0(\mathbf{A})^{-1}$ exists, $\forall\, \mathbf{A} \in V$; see Problems 9 and 10 to §5.10, p.477. It follows that for all $\mathbf{A}$'s outside a set of zero Lebesgue measure, there is a number C, depending on $\mathbf{A}$, such that the above inequality holds. The question of determining or estimating a number C such that $C > \sup_{V 0 \ne \boldsymbol{\nu} \in \mathcal{Z}^\ell} |\boldsymbol{\nu} \cdot \boldsymbol{\omega}|^{-1} |\boldsymbol{\nu}|^{-\ell}$ is, for a given $\boldsymbol{\omega}$, an interesting and difficult number-theoretic problem. Some of its aspects are discussed in detail in the problems of §2.20.

(4) Suppose $f_0 = \lambda \widetilde{f}_0$, with $\widetilde{f}_0$ λ-independent, and fix a set $\widetilde{V} \subset V$, bounded and closed. Using the notations of the Observation (2) let

$$\varrho_0 = \min_{\mathbf{A}\in V} \varrho(\mathbf{A}), \qquad \xi_0 = \min_{\mathbf{A}\in V} \xi(\mathbf{A}),$$
$$E_0 = \max_{\mathbf{A}\in V} |\frac{\partial h_0}{\partial \mathbf{A}}|_{\varrho(\mathbf{A})}, \qquad \eta_0 = \max_{\mathbf{A}\in V} ||M_0^{-1}||_{\varrho(\mathbf{A})}.$$

Then apply Proposition 22 to the Hamiltonian system described by Eq. (5.12.3) in $\mathcal{S}_{\varrho_0}(\mathbf{A}_0) \times \mathcal{T}^\ell$ with $\mathbf{A}_0 \in \widetilde{V}(C)$, see Observation (3) above. By Eq. (5.12.6), one immediately deduces that for λ small, the perturbed Hamiltonian system admits simultaneously coexisting invariant tori $\mathcal{T}(\boldsymbol{\omega}(\mathbf{A}_0)), \forall\, \mathbf{A}_0 \in \widetilde{V}(C)$. Such tori will be located geometrically close to the unperturbed tori, by Eq. (5.12.9).
This means that the "less resonant" the pulsations $\boldsymbol{\omega}$ of the unperturbed quasi-periodic motions,[31] the larger the perturbations intensity λ has to become before it can possibly succeed in destroying these motions and the invariant tori on which they take place.

(5) Observations (1)-(4) above show that the statement (i) of Proposition 15, p.460, §5.9, and the statement that $W^{(\varepsilon)} \neq \emptyset$ follow from Proposition 22. From the proof of Proposition 22, however, all of Proposition 15 follows with some effort, in the analytic case. We shall not discuss this problem, (see for instance [39], [12], [21]).

(6) The condition (5.12.6) involves only the derivatives of h_0 and f_0: this is natural since only such functions appear in the equations of motion.
Also, it should be noted that the nature of the condition (5.12.6) is quite simple: given $h_0, f_0, \mathbf{A}_0$, one can form the quantities E_0, ε_0, C and, with them, the "dimensionless quantities" $C\varepsilon_0, CE_0$, $\eta_0\varrho_0^{-1}E_0, \xi_0$ in terms of which all the other dimensionless quantities can be formed. It can be seen that

$$CE_0 \geq 1, \qquad \eta_0\varrho_0^{-1}E_0 \geq 1, \tag{5.12.11}$$

see Problem 1 at the end of the section. Then Eq. (5.12.6) just says that the perturbation strength $C\varepsilon_0$, has to be small compared to the other "small" dimensionless quantities $(CE_0)^{-1}$, $(\eta_0\varrho_0^{-1}E_0)^{-1}$ and ξ_0 which are relevant to the problem.
Note that in the above argument, the parameters "relevant to the problem" are just $E_0, \varepsilon_0, C, \varrho_0, \xi_0, \eta_0$: this is, in fact, not obvious and, a priori, one might expect that other quantities may be relevant, like $F_0 = |\frac{\partial^2 h_0}{\partial \mathbf{A}^2}|_{\varrho_0}$ or $\widetilde{F}_0 = |\frac{\partial^3 h_0}{\partial \mathbf{A}^3}|_{\varrho_0}$, etc. All that the above argument says is that if the results of Proposition 22 hold under conditions that just involve $E_0, \varepsilon_0, C, \varrho_0, \xi_0, \eta_0$, then it is not surprising that such conditions can take the form of Eq. (5.12.6), i.e., the simplest imaginable form.

(7) The condition $\eta_0 < +\infty$ or something like it must be necessary: in fact, for isochronous systems the above theorem cannot hold. Just consider $\ell =$

[31] i.e., the smaller C is.

1, $h_0(A) = A$, $f_0(A,\varphi) = \varepsilon A$; in this case all the perturbed motions have pulsations $\omega = 1+\varepsilon$ and none $= \omega_0 = 1$. The parameter η_0 will be called the "anisochrony parameter" and a system for which $\eta_0 < +\infty$ is said to be "anisochronous" near $\mathbf{A}_0$.

The systems of harmonic oscillators are strictly isochronous, and the theorem does not directly apply to them.

However, if $h_0(\mathbf{A}) = \boldsymbol{\omega}_0\cdot\mathbf{A}$ and if $\boldsymbol{\omega}_0$ verifies Eq. (5.12.4), then the theorem can still be indirectly applied under some additional assumptions. In fact, let $f_0 = \lambda\widetilde{f}_0$ with $\widetilde{f}_0$ λ-independent. Assume that

$$\widetilde{\eta}_0 = ||(\frac{\partial\widetilde{f}_{0\mathbf{0}}}{\partial\mathbf{A}\partial\mathbf{A}})^{-1}|| < +\infty \tag{5.12.12}$$

where $\widetilde{f}_{0\mathbf{0}}$ denotes the average of f_0 over $\mathcal{T}^\ell$, i.e., its Fourier coefficient with $\boldsymbol{\nu} = \mathbf{0}$ [see also Eq. (5.11.16)]. Now apply Proposition 17, p.469, to change variables completely canonically and to transform the problem into that of the analysis of the systems with Hamiltonian

$$h_0'(\mathbf{A}) + \lambda^{(n+1)} f_0'(\mathbf{A},\boldsymbol{\varphi}), \tag{5.12.13}$$

where h_0', f_0' are holomorphic in $C(\frac{1}{2}\varrho_0, \frac{1}{2}\xi_0; \mathbf{A}_0)$ and in the variable $\mathbf{A}$ for λ close to zero, see Eq. (5.10.27); choose n as $n = a\ell + b + \ell$, a and b being the constant in Eq. (5.12.6). Also, from Eq. (5.10.41), we see that

$$h_0'(\mathbf{A}) = h_0(\mathbf{A}) + \lambda\widetilde{f}_{0\mathbf{0}}(\mathbf{A}) + \lambda^2\widetilde{h}(\mathbf{A}), \tag{5.12.14}$$

where $\widetilde{h}$ is analytic in λ (actually, it is a polynomial) and in $\mathbf{A}$, near $\mathbf{A}_0$.

Therefore, if λ is small enough, the quantities $E_0', \eta_0', \varepsilon_0'$ such that

$$\begin{aligned} E_0' \geq& |\frac{\partial h_0'}{\partial\mathbf{A}}|_{\varrho_0/2}, \qquad \eta_0' \geq ||(\frac{\partial^2 h_0'}{\partial\mathbf{A}\partial\mathbf{A}})^{-1}||_{\varrho_0/2},\\ \varepsilon_0' \geq& \lambda^{b\ell+1}\Big(|\frac{\partial f_0'}{\partial\mathbf{A}}|_{\varrho_0/2,\,\xi_0/2} + \frac{2}{\varrho_0}|\frac{\partial f_0'}{\partial\boldsymbol{\varphi}}|_{\varrho_0/2,\,\xi_0/2} + \frac{2}{\varrho_0}\Big) \end{aligned} \tag{5.12.15}$$

can be chosen so that for a suitable $K > 0$, depending on $E_0, \varepsilon_0, \varrho_0, \xi_0$ but not on λ, and $\forall\,\lambda$ small:

$$E_0' \leq 2E_0, \quad \eta_0' \leq 2\widetilde{\eta}_0\lambda^{-1}, \quad \varepsilon_0' \leq K\lambda^{b\ell+1}. \tag{5.12.16}$$

Consider, next, the points $\mathbf{A}_0' \in \mathcal{S}_{\frac{1}{2}\varrho_0}(\mathbf{A}_0)$ with $\boldsymbol{\omega}'(\mathbf{A}_0') = \frac{\partial h_0'}{\partial\mathbf{A}}(\mathbf{A}_0')$ such that

$$|\boldsymbol{\omega}'(\mathbf{A}_0')\cdot\boldsymbol{\nu}|^{-1} \leq C\lambda^{-\ell}|\boldsymbol{\nu}|^\ell, \qquad \forall\,\boldsymbol{\nu}\in\mathcal{Z}_+^\ell,\ |\boldsymbol{\nu}| > 0 \tag{5.12.17}$$

Using the results of the Problems 9 and 15, §5.10, p.477 and 478, and the estimate on η_0', it is possible to see that such points actually exist and fill a considerable part of $\mathcal{S}_{\frac{1}{4}\varrho_0}(\mathbf{A}_0')$ (in fact, their ensemble forms a set whose measure approaches that of $\mathcal{S}_{\frac{1}{4}\varrho_0}(\mathbf{A}_0')$ itself as $C\to\infty$, uniformly in λ).

Proposition 20 can be applied to $h_0' + \lambda^{b\lambda+\ell+1} f_0'$, regarded as holomorphic on $C(\frac{1}{4}\varrho_0, \frac{1}{4}\xi_0; \mathbf{A}_0)$ with $\mathbf{A}_0'$ verifying Eq. (5.12.17), and Eq. (5.12.6) becomes

$$BC\lambda^{-\ell} K \lambda^{a\ell+b\ell+1} (2\lambda^{-1} E_0)^a (2\widetilde{\eta}_0 \lambda^{-1} 4\varrho_0^{-1} 2E_0)^b (4\xi_0^{-1})^c < 1$$

which can be fulfilled for λ small.
This could be interpreted as saying that the quasi-periodic motions with $\mathbf{A}_0'$ such that Eq. (5.12.17) holds are not destroyed by the perturbation, but survive with a slightly modified pulsation (since $\boldsymbol{\omega}'(\mathbf{A}_0') = \boldsymbol{\omega}_0 + O(\lambda)$), running on slightly deformed tori.

(8) So Observation (7) shows that the non isochrony condition, $\eta_0 < +\infty$, can be essentially weakened. One can ask whether this is the case for the "non resonance" condition $C < +\infty$ as well. The whole discussion of perturbation theory, §5.10, suggests that this is not the case.
In fact, by considering some extreme cases, it appears that one cannot go too far toward weakening the conditions. Consider a harmonic isochronous resonating oscillators in $\mathcal{R}^3$:

$$H(\mathbf{p}, \mathbf{q}) = \frac{1}{2}(\mathbf{p}^2 + \mathbf{q}^2). \tag{5.12.18}$$

and use action-angle coordinates $(\mathbf{A}, \boldsymbol{\varphi}) \in (\mathcal{R}_+^{/}0)^3 3 \times \mathcal{T}^3$ to describe (most of) the motions via the Hamiltonian[32]

$$h_0(\mathbf{A}) = A_1 + A_2 + A_3$$

on $(0, +\infty)^3 \times \mathcal{T}^3$. A further completely canonical change of coordinates, $\mathbf{A} \to \widetilde{\mathbf{A}}$, $\boldsymbol{\varphi} \to \widetilde{\boldsymbol{\varphi}}$:

$$\begin{aligned} \widetilde{A}_1 =& A_1 + A_2 + A_3, & \widetilde{\varphi}_1 &= \varphi_1, \\ \widetilde{A}_2 =& A_2, & \widetilde{\varphi}_2 &= \varphi_2 - \varphi_1, \\ \widetilde{A}_3 =& A_3, & \widetilde{\varphi}_3 &= \varphi_3 - \varphi_1, \end{aligned} \tag{5.12.19}$$

(see Problem 33, §3.11, p.232) transforms the Hamiltonian (5.12.18) into

$$\widetilde{h}_0(\widetilde{\mathbf{A}}) = \widetilde{A}_1, \quad (\widetilde{\mathbf{A}}, \widetilde{\boldsymbol{\varphi}}) \in (0, +\infty)^3 \times \mathcal{T}^3. \tag{5.12.20}$$

Let $f_0(\widetilde{A}_2, \widetilde{A}_3, \widetilde{\varphi}_2, \widetilde{\varphi}_3)$ be an analytic non integrable Hamiltonian on $V \times \mathcal{T}^2 \subset (0, +\infty)^2 \times \mathcal{T}^2$: its existence is not obvious, but we state without proof that it exists and that it can be chosen so that it produces non quasi periodic motions.[33] Then the system

$$\widetilde{h}_0(\widetilde{\mathbf{A}}) + \varepsilon f_0(\widetilde{A}_2, \widetilde{A}_, \widetilde{\varphi}_2, \widetilde{\varphi}_3) \tag{5.12.21}$$

[32] See exercises for §4.1.
[33] An example could be constructed on the basis of Observation (3), p.336, but the discussion is quite long.

cannot be integrable as, manifestly, the coordinates corresponding to the degrees of freedom with indices 2 and 3 verify the equations with Hamiltonian εf_0 which gives rise, for $\varepsilon \neq 0$, to motions coinciding with those of f_0, up to a change of scale in time and which are not quasi periodic, i.e., not integrable by criterion (i), p.353.
The example shows why resonances can be important. In a resonant situation, it happens that some degrees of freedom of the system "do not move at all" as can be seen by suitable changes of coordinates. Hence, upon perturbation, their motion will be entirely governed by the perturbation and it will therefore become important whether or not the perturbation by itself is integrable.
If the perturbation by itself describes an integrable system in the phase space region around a resonant torus of the unperturbed system, the above argument suggests that something could, nevertheless, be done. This is in fact the situation found in celestial mechanics in the vicinity of the unperturbed tori corresponding to orbits of small eccentricity and small inclination. As shown in §5.10, Observation (3), p.473, in this situation one can set up some perturbation scheme to compute the secular perturbations. The scheme can lead to a rigorous proof of tori conservation (under suitable assumptions on the phase-space region which is considered). This proof is in a celebrated paper by Arnold, [3].
Of course, in the above discussion, one could have directly started from Eqs. (5.12.20) and (5.12.21), but we thought that starting from a physical system would be easier for the reader. On the other hand, the choice of $\mathcal{R}^3$ is essential to the argument: if we had chosen $\mathcal{R}^2$, the argument could have failed since only $\widetilde{A}_2, \widetilde{\varphi}_2$ would have been present, i.e., f_0 would have described a one-degree-of-freedom system (which is "necessarily"[34] integrable).

(9) The above observation shows that the non resonance condition is essential in a case in which the resonance is very manifest, i.e., the unperturbed system is isochronous and resonating. However, one could think that the non integrability phenomenon might be only related to isochronous resonances: if a system is anisochronous one might argue that the perturbation will cause the motion to wander around in phase space, keeping it away from the resonances most of the time. The fallaciousness of this way of reasoning is made clear by an example that goes back to Poincaré. Consider the system on $\mathcal{R}^2 \times \mathcal{T}^2$:

$$H(A_1, A_2, \varphi_1, \varphi_2) = \frac{1}{2}(A_1^2 + A_2^2) + \varepsilon\, f(\varphi_1, \varphi_2), \qquad \text{with} \tag{5.12.22}$$

$$g(\varphi_2 - \varphi_1) \stackrel{def}{=} \int_0^{2\pi} f(\varphi_1 + \psi, \varphi_2 + \psi) \frac{d\psi}{2\pi} = \text{not constant} \tag{5.12.23}$$

To fix the ideas we shall take $f(\varphi_1, \varphi_2) = 1 - \cos(\varphi_2 - \varphi_1)$: in this case Eq. (5.12.22) has a simple physical meaning, as it describes two points ideally

[34] See the statement (19), p.363, and §2.7 for general conditions of integrability.

bound to a unit circle attracting each other via a harmonic force. The reader should, as an exercise, understand the physical meaning of the argument below and why it can be immediately extended to the general case if (5.12.23) holds. For $\varepsilon = 0$ all the motions on the torus $\{\mathbf{A}_0\} \times \mathcal{T}^2$, $\mathbf{A}_0 = (1,1)$, are periodic with pulsations $\boldsymbol{\omega}_0 = \mathbf{A}_0 = (1,1)$, so the torus is resonant.

Suppose, per absurdum, that the torus is not destroyed for small ε, in the sense that there exists an invariant torus (i.e., invariant with respect to the perturbed motion) with parametric equations:

$$\begin{aligned} \mathbf{A} =&\mathbf{A}_0 + \boldsymbol{\alpha}(\boldsymbol{\varphi}'), \\ \boldsymbol{\varphi} =&\boldsymbol{\varphi}' + \boldsymbol{\alpha}(\boldsymbol{\varphi}'), \end{aligned} \tag{5.12.24}$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are $\mathcal{R}^2$-valued functions in $C^\infty(\mathcal{T}^2)$, and that the torus given by Eq. (5.12.24) is close to the unperturbed torus for ε small:

$$\gamma(\varepsilon) = \max|\boldsymbol{\alpha}(\boldsymbol{\varphi}')|\varrho_0^{-1} + \max|\boldsymbol{\beta}(\boldsymbol{\varphi}')| \xrightarrow[\varepsilon\to 0]{} 0. \tag{5.12.25}$$

Suppose also that the motion on the torus in Eq. (5.12.24) is described by $\boldsymbol{\varphi}' \to \boldsymbol{\varphi}'' + \boldsymbol{\omega}\, t$, $\boldsymbol{\omega} = (1,1) = \mathbf{A}_0$, i.e., assume that the perturbed torus is run periodically with the same spectrum as that corresponding to the unperturbed torus $\{\mathbf{A}_0\} \times \mathcal{T}^2$. Write the Hamiltonian equations for $h_0 + \varepsilon f$ and subtract the two equations for A_1 and A_2:

$$\dot{A}_2 - \dot{A}_1 = -2\varepsilon \sin(\varphi_2 - \varphi_1). \tag{5.12.26}$$

Then integrate both sides between $t = 0$ and $t = 2\pi$, assuming that to have computed them on a motion developing on the torus of Eq. (5.12.24) with initial datum corresponding to $\boldsymbol{\varphi}' \in \mathcal{T}^2$. Since the motion is periodic, by assumption, with period 2π ($\boldsymbol{\omega}_0 = (1,1)$), it is

$$\begin{aligned} 0 = -2\varepsilon &\int_0^{2\pi} \sin(\varphi_2 - \varphi_1) dt \qquad (5.12.27)\\ =&- 2\varepsilon \int_0^{2\pi} \sin[\varphi_2' - \varphi_1' + \beta_2((\varphi_1' + t, \varphi_2' + t)) - \beta_1((\varphi_1' + t, \varphi_2' + t))]\, dt \\ =&- 4\pi\varepsilon([\sin(\varphi_2' - \varphi'1)] + 2\widetilde{\gamma}(\varepsilon)) \xrightarrow[\varepsilon\to 0]{} \sim -4\pi\varepsilon \sin(\varphi_2' - \varphi_2'), \end{aligned}$$

where $\widetilde{\gamma}(\varepsilon) \in [-\gamma(\varepsilon), \gamma(\varepsilon)]$ is suitably chosen. This is absurd if $\varphi_2' - \varphi_1' \neq 0, \pi$ and shows that the torus of Eq. (5.12.24) cannot exist as an invariant torus run periodically with pulsation $\boldsymbol{\omega}_0 = (1,1)$. The resonating torus corresponding to $\mathbf{A}_0 = (1,1)$ is "destroyed" upon perturbation, no matter how small.

The argument shows that the torus is "destroyed", but does not show that all the periodic motions with period 2π are destroyed. For instance if $\varphi_1 = \varphi_2$ or $\varphi_1 = \varphi_2 + \pi$ we form, together with $A_1 = A_2 = 1$, two sets of initial data evolving periodically with period 2π and, topologically, such sets are two circles (i.e., like $\mathcal{T}^1$ instead of $\mathcal{T}^2$).

This example is interesting because it considers a case in which all the assumptions of Proposition 22 hold except the non resonance condition (5.12.4), thereby showing its necessity. However, it does not provide an example as "shocking" as the one of observation (8), since the perturbed system still exhibits only quasi-periodic motions or motions with rather trivial asymptotic behavior.[35] Much more interesting in this respect would be the case when f in Eq. (5.12.23) is replaced by a function really depending on both φ_1, and φ_2, not only on $\varphi_2 - \varphi_1$. In such a case, one expects to find some motions with very complex asymptotic behavior near a resonating unperturbed torus.

(10) Observations (7)-(9) above clarify the necessity of the assumptions in Proposition 22. They can be summarized as follows: non resonating quasi-periodic motions on ℓ-dimensional tori are preserved, in anisochronous systems, in the presence of small perturbations; they are also preserved in isochronous non resonating systems for all the non isochronous small perturbations (modulo a small change in the frequencies). Resonating motions on ℓ-dimensional tori are generally destroyed by small perturbations in both the isochronous and the non isochronous cases.

(11) It is important that the reader who is about to read the following proof realizes that all the very numerous inequalities that will be met can easily be guessed on "dimensional grounds", i.e., using what we called in §5.11, Observation (2), p.483, "dimensional estimates". In this way, one can easily check the calculations (which we give in great detail only for completeness since this book is supposed to be elementary).
The possibility of simple dimensional estimates is what makes the proof in the analytic case easy to visualize.
In the upcoming proof no attention is paid to optimal estimates, nor to the evaluation of the various constants. However, in principle, the proof below does not contain any crude approximation, and if the constants are evaluated with care it should give results which are optimal *in the given generality of the assumptions.*
This, of course, does not mean that in particular cases the estimates could not be greatly improved.
Finally let us point out to the reader familiar with present trends in statistical mechanics and field theory that the proof below yields a nice example of a vast class of theorems which can be proved by what has become known in physics as the "renormalization group method".

Proof. We think of the unperturbed Hamiltonian h_0 and the perturbation f_0 as a pair of holomorphic functions on $C(\varrho_0, \xi_0; \mathbf{A}_0) \subset \mathcal{C}^{2\ell}$, real on $\mathcal{S}_{\varrho_0} \times \mathcal{T}^\ell$, see Definition 9, p.481 487, §5.11. To this pair we associate the "characteristic numbers" $E_0, \eta_0, \varrho_0, \xi_0, \varepsilon_0$ verifying Eq. (5.12.5).

We have already noted that [see Eq. (5.12.11)]

[35] as can be seen by the completely canonical change of variables $A = \frac{1}{2}(A_1 + A_2)$, $B = \frac{1}{2}(A_1 - A_2 z)$, $\varphi = \varphi_1 + \varphi_2, \psi = \varphi_1 - \varphi_2$, (exercise).

$$\eta_0 \varrho_0^{-1} E_0 \geq 1, \qquad C E_0 \geq 1. \tag{5.12.28}$$

In the course of the proof, we shall have to "give up" some analyticity in the $\mathbf{A}$ and $\boldsymbol{\varphi}$ variables in order to make dimensional estimates. The amount of analyticity that is given up is, to a great extent, arbitrary: we introduce some "analyticity loss" parameters $\delta_0 > \delta_1 > \ldots$ which will be used to describe precisely the analyticity loss. To be definite, let

$$\delta_k = \frac{1}{2^4} \frac{\xi_0}{(1+k)^2} \tag{5.12.29}$$

so that $5 \sum_{k=0}^{\infty} \delta_k < \xi_0 < 1$. For simplicity, assume that $C\varepsilon_0 E < 1$.

The identification of $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_\ell) \in \mathcal{T}^\ell$ with $\mathbf{z} = (e^{i\varphi_1}, \ldots, e^{i\varphi_\ell}) \in \mathcal{C}^\ell$ will be often used, while also freely using an "angular notation" for $\mathbf{z}$ even if $\mathbf{z}$ is not on the product of the ℓ unit circles. In this case, $\partial/\partial\varphi_k$ means $iz_k\partial/\partial z_k$, see Definition 9 (ii), p.481. Also, it will be convenient to write $e^{i\boldsymbol{\Delta}} \equiv (e^{i\Delta_1}, \ldots, e^{i\Delta_\ell})$ and $\mathbf{z}\, e^{i\boldsymbol{\Delta}} \equiv (z_1 e^{i\Delta_1}, \ldots, z_\ell e^{i\Delta_\ell})$ for $\boldsymbol{\Delta} \in \mathcal{C}^\ell$. Such conventions greatly simplify the notations.

The proof proceeds by applying perturbation theory along the lines of §5.10. Since the first problem is that f_0 does not fulfill the assumptions of Proposition 17, we shall divide f_0 into two parts: one very small $O(\varepsilon_0^2)$ and the other fulfilling the assumptions of Proposition 17, i.e., with only finitely many Fourier components ("Arnold regularization").

Then we shall apply Proposition 17 to find a canonical transformation changing the Hamiltonian into a "renormalized" one with an integrable part $h_1(\mathbf{A})$ plus a perturbation $f_1(\mathbf{A}, \boldsymbol{\varphi})$ with f_1 of $O(\varepsilon_0^2)$). Afterwards, we proceed to find a point $\mathbf{A}_1$ such that $\frac{\partial h_1}{\partial \mathbf{A}}(\mathbf{A}_1) = \boldsymbol{\omega}_0$, and we shall again be in a position to begin the procedure all over again, provided we control the new characteristic parameters $E_1, \varepsilon_1, \varrho_1, \xi_1, \eta_1$. Basically, the whole argument is reduced to searching for an expression of $E_1, \varepsilon_1, \varrho_1, \xi_1, \eta_1$ in terms of $E_0, \varepsilon_0, \varrho_0, \xi_0, \eta_0$ ("Kolmogorov's iteration").

To reduce f_0 to a trigonometric polynomial plus a small remainder, introduce the "ultraviolet cut off":

$$N_0 = \frac{2}{\delta_0} \log \frac{1}{C\varepsilon_0 \delta_0^\ell} > 1 \tag{5.12.30}$$

and define the "regularized perturbation"

$$f_0^{(\leq N_0]}(\mathbf{A}, \boldsymbol{\varphi}) \stackrel{def}{=} \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell,\, |\boldsymbol{\nu}| \leq N_0} f_{0\boldsymbol{\nu}}(\mathbf{A})\, e^{i\,\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}, \tag{5.12.31}$$

using for f_0 the notation of Eqs. (5.11.16), p.482. Let $f^{[>N_0]} \stackrel{def}{=} f_0 - f^{(\leq N_0]}$.

The choice of N_0 has been made so that $f_0^{[>N_0]}$ is indeed of $O(\varepsilon_0^2)$. This can be seen by applying the estimates of Eqs. (5.11.15) to the functions $\frac{\partial f}{\partial \mathbf{A}}$

and $\varrho_0^{-1}\frac{\partial f_0}{\partial \boldsymbol{\varphi}}$, holomorphic in $C(\varrho_0, \xi_0; \mathbf{A}_0)$, regarded as function on $C(\xi)$ parameterized by $\mathbf{A} \in \widehat{\mathcal{S}}_{\varrho_0(\mathbf{A}_0)}$, $\forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell$, $\forall\, i = 1, \ldots, \ell$,

$$|\frac{\partial f_{0\boldsymbol{\nu}}}{\partial \mathbf{A}}| \le \varepsilon_0\, e^{-\xi_0|\boldsymbol{\nu}|}, \qquad |\nu_i f_{0\boldsymbol{\nu}}| \le \varepsilon_0\, \varrho_0\, e^{-\xi_0|\boldsymbol{\nu}|}, \tag{5.12.32}$$

by the third of Eqs. (5.11.15). Also note that by item (v) Proposition 18, p.481, the functions $\frac{\partial f_{0\boldsymbol{\nu}}(\mathbf{A})}{\partial \mathbf{A}}$ and $\nu_i\, f_{0\boldsymbol{\nu}}(\mathbf{A})$ are $\forall\, \boldsymbol{\nu} \in \mathcal{Z}^\ell$, $\forall\, i = 1, \ldots, \ell$, holomorphic on $\widehat{\mathcal{S}}_{\varrho_0}(\mathbf{A}_0)$. Just apply (v) to the functions $g = \frac{\partial f_0}{\partial \mathbf{A}}$ and $g = \frac{\partial f_0}{\partial \boldsymbol{\varphi}}$.

Equation (5.12.32) allow us to bound $f_0^{[>N_0]}$ and $f_0^{[\le N_0]}$ as follows. There exist B_1, B_2, $1 \le< B_1 \le B_2$ such that

$$\begin{aligned} &|\frac{\partial f_{0\boldsymbol{\nu}}^{[\le N_0]}}{\partial \mathbf{A}}|_{\varrho_0, \xi_0-\delta_0} + \frac{1}{\varrho_0}|\frac{\partial f_{0\boldsymbol{\nu}}^{[\le N_0]}}{\partial \boldsymbol{\varphi}}|_{\varrho_0, \xi_0-\delta_0} \le B_1 \varepsilon_0 \delta_0^{-1}, \\ &|\frac{\partial f_{0\boldsymbol{\nu}}^{[>N_0]}}{\partial \mathbf{A}}|_{\varrho_0, \xi_0-\delta_0} + \frac{1}{\varrho_0}|\frac{\partial f_{0\boldsymbol{\nu}}^{[>N_0]}}{\partial \boldsymbol{\varphi}}|_{\varrho_0, \xi_0-\delta_0} \le B_2 \varepsilon_0^2 C. \end{aligned} \tag{5.12.33}$$

These estimates follow by substituting the bounds given by Eq. (5.12.32) into Eq. (5.12.31) or into the analogous expression for $f_0^{[>N_0]}$, after the appropriate differentiations. For instance, consider the second of Eqs. (5.12.33). One has, $\forall\, (\mathbf{A}, \mathbf{z}) \in C(\varrho_0, \xi_0 - \delta_0; \mathbf{A}_0)$,

$$\begin{aligned} &|\frac{\partial f_{0\boldsymbol{\nu}}^{[>N_0]}(\mathbf{A}, \boldsymbol{\varphi})}{\partial \mathbf{A}}| = |\sum_{\substack{\boldsymbol{\nu} \in \mathcal{Z}^\ell \\ |\boldsymbol{\nu}| > N_0}} \frac{\partial f_{0\boldsymbol{\nu}}(\mathbf{A})}{\partial \mathbf{A}} e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}| \equiv |\sum_{\substack{\boldsymbol{\nu} \in \mathcal{Z}^\ell, \\ |\boldsymbol{\nu}| > N_0}} \frac{\partial f_{0\boldsymbol{\nu}}(\mathbf{A})}{\partial \mathbf{A}} \mathbf{z}^{\boldsymbol{\nu}}| \\ &\le \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell,\, |\boldsymbol{\nu}| > N_0} \varepsilon_0\, e^{-\delta_0|\boldsymbol{\nu}|} \le \varepsilon_0\, e^{-\frac{\delta_0}{2} N_0} \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell,\, |\boldsymbol{\nu}| > N_0} \varepsilon_0\, e^{-\frac{1}{2}\delta_0|\boldsymbol{\nu}|} \\ &= C\varepsilon_0^2 \delta_0^\ell \big(\frac{1 + e^{-\frac{1}{2}\delta_0}}{1 + e^{+\frac{1}{2}\delta_0}}\big)^\ell \le B' C \varepsilon_0^2, cr \end{aligned} \tag{5.12.34}$$

where in the first equality, we use the symbolic but suggestive "angular notation" for $\mathbf{z}$, and $B' > 0$ is a suitable constant. Similarly,

$$|\frac{\partial f_{0\boldsymbol{\nu}}^{[>N_0]}(\mathbf{A}, \mathbf{z})}{\partial \boldsymbol{\varphi}}| \equiv |\sum_{\substack{\boldsymbol{\nu} \in \mathcal{Z}^\ell \\ |\boldsymbol{\nu}| > N_0}} \boldsymbol{\nu}\, f_{0\boldsymbol{\nu}}(\mathbf{A}) \mathbf{z}^{\boldsymbol{\nu}}| \le \varepsilon_0 \varrho_0 \sum_{|\boldsymbol{\nu}| > N_0} e^{-\delta_0|\boldsymbol{\nu}|} \le B' C \varepsilon_0^2 \varrho_0. \tag{5.12.35}$$

Hence, the second of Eqs. (5.12.33) follows from Eqs. (5.12.34) and (5.12.35). The first of Eqs. (5.12.33) follows from the same type of arguments.[36]

---

[36] One could take, say, $B_1 = B_2 = 2\,(4\sqrt{e})^\ell$, because $\frac{1+e^{-\frac{1}{2}\delta_0}}{1+e^{-\frac{1}{2}\delta_0}} < \frac{4\sqrt{e}}{\delta_0}$, if $\delta_0 < 1$.

Following the ideas of perturbation theory, a canonical change of variables will be constructed using, as in the proof to Proposition 16, §5.10, p.466, a generating function of the form $(\mathbf{A}', \boldsymbol{\varphi}) \to \mathbf{A}' \cdot \boldsymbol{\varphi} + \Phi_0(\mathbf{A}', \boldsymbol{\varphi})$, where $\boldsymbol{\varphi}_0$ is defined on a suitable set $\mathcal{S}_{\widetilde{\varrho}_0} \times \mathcal{T}^\ell$ as

$$\Phi_0(\mathbf{A}', \boldsymbol{\varphi}) = \sum_{\boldsymbol{\nu} \in \mathcal{Z}^\ell,\, 0<|\boldsymbol{\nu}|\le N_0} \frac{f_{0\boldsymbol{\nu}}(\mathbf{A}')\, \mathbf{z}^{\boldsymbol{\nu}}}{-i\boldsymbol{\omega}(\mathbf{A}') \cdot \boldsymbol{\nu}} \tag{5.12.36}$$

which defines a holomorphic function of $(\mathbf{A}', \mathbf{z}) \in C(\widetilde{\varrho}_0, \xi_0 - \delta_0; \mathbf{A}_0)$ if $\widetilde{\varrho}_0$ is chosen so small that, by consequence of Eq. (5.12.4), $|\boldsymbol{\omega}(\mathbf{A}) \cdot \boldsymbol{\nu}| > 0\ \forall \boldsymbol{\nu} \in \mathcal{Z}^\ell, \boldsymbol{\nu} \neq \mathbf{0}$, $\forall\, \mathbf{A} \in \widehat{\mathcal{S}}_{\widetilde{\varrho}_0}(\mathbf{A}_0)$. Actually, a simple choice for $\widetilde{\varrho}_0$, good enough for our purposes. In fact, $\forall\, \mathbf{A}' \in \widehat{\mathcal{S}}_{\widetilde{\varrho}_0}(\mathbf{A}_0)$ and if $\widetilde{\varrho}_0 < \frac{1}{2}\varrho_0$, it is

$$\begin{aligned} |\boldsymbol{\omega}(\mathbf{A}') \cdot \boldsymbol{\nu}|^{-1} &\equiv |(\boldsymbol{\omega}_0 + (\boldsymbol{\omega}(\mathbf{A}') - \boldsymbol{\omega}(\mathbf{A}_0))) \cdot \boldsymbol{\nu}|^{-1} \\ &\le |\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}|^{-1} \Big|1 - \frac{|(\boldsymbol{\omega}(\mathbf{A}') - \boldsymbol{\omega}(\mathbf{A}_0)) \cdot \boldsymbol{\nu}|}{|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}|}\Big|^{-1} \\ &\le C\, |\boldsymbol{\nu}|^\ell\, \big|1 - 2C|\boldsymbol{\nu}|^{\ell+1} \ell E_0 \frac{\widetilde{\varrho}_0}{\varrho_0}\big|^{-1} \end{aligned} \tag{5.12.37}$$

because we can bound $|\boldsymbol{\omega}(\mathbf{A}') - \boldsymbol{\omega}(\mathbf{A}_0)|$ as

$$\begin{aligned} |\boldsymbol{\omega}(\mathbf{A}') - \boldsymbol{\omega}(\mathbf{A}_0)| &\equiv \Big| \int_0^1 dt\, \frac{d}{dt} \boldsymbol{\omega}(\mathbf{A}_0 + t\,(\mathbf{A} - \mathbf{A}_0)) \Big| \\ =& \Big| \int_0^1 dt \Big(\sum_{j=1}^{\ell} \frac{\partial \boldsymbol{\omega}(\mathbf{A}_0 + t\,(\mathbf{A}' - \mathbf{A}_0))}{\partial A_j}\, (A'_j - A_{0j}) \Big)\, dt \\ \le& \ell \frac{E_0}{\varrho_0 - \widetilde{\varrho}_0}\, \widetilde{\varrho}_0 \le 2\ell E_0 \frac{\widetilde{\varrho}_0}{\varrho_0} \end{aligned} \tag{5.12.38}$$

by a dimensional estimate like the first of Eqs. (5.11.18); hence, if

$$\widetilde{\varrho}_0 \stackrel{def}{=} \frac{\varrho_0}{2\ell C E_0 N_0^{\ell+1}}, \tag{5.12.39}$$

then, since $CE_0 \ge 1, N_0 \ge 1, \widetilde{\varrho}_0 < \frac{1}{2}\varrho_0$, that Eq. (5.12.37) implies

$$|\boldsymbol{\omega}(\mathbf{A}') \cdot \boldsymbol{\nu}|^{-1} < 2C|\boldsymbol{\nu}|^\ell, \qquad \forall\ 0 < |\boldsymbol{\nu}| \le N_0 \tag{5.12.40}$$

for $\mathbf{A}' \in \widehat{\mathcal{S}}_{\varrho_0(\mathbf{A}_0)}$. Hence, Eq. (5.12.36) implies that Φ_0 is holomorphic in $C(\widetilde{\varrho}_0, \xi_0 - \delta_0; \mathbf{A}_0)$ and that, using the second of Eqs. (5.12.32),[37]

[37] Recall that $\boldsymbol{\nu}|, \boldsymbol{\nu} \in \mathcal{Z}^\ell$ and $|\mathbf{w}|, \mathbf{w} \in \mathcal{C}^\ell$, have a different meaning by our conventions, Eqs. (5.11.1) and (5.11.2). This explains the factor ℓ.

$$\begin{aligned} |\Phi_0(\mathbf{A}',\mathbf{z})| \leq & \sum_{0<|\boldsymbol{\nu}|\leq N_0} \frac{|f_{0\boldsymbol{\nu}}(\mathbf{A}')|}{|\boldsymbol{\omega}(\mathbf{A}')\cdot\boldsymbol{\nu}|} e^{(\xi_0-\delta_0)|\boldsymbol{\nu}|} \\ \leq & \sum_{0<|\boldsymbol{\nu}|\leq N_0} 2C|\boldsymbol{\nu}|^{\ell}|f_{0\boldsymbol{\nu}}(\mathbf{A}')|e^{(\xi_0-\delta_0)|\boldsymbol{\nu}|} \\ \leq & \sum_{0<|\boldsymbol{\nu}|} 2C|\boldsymbol{\nu}|^{\ell-1}\lambda\varepsilon_0\varrho_0 e^{-\delta_0|\boldsymbol{\nu}} \leq B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell+1} \end{aligned} \tag{5.12.41}$$

for all $(\mathbf{A}',\mathbf{z}) \in C(\wr\varrho_0, \xi_0-\delta_0; \mathbf{A}_0)$, with $B_3 > 2$.[38]

Hence, by the dimensional estimates of Eq. (5.11.18),

$$\begin{aligned} \Big|\frac{\partial\Phi_0}{\partial\mathbf{A}'}\Big|_{\widetilde{\varrho}_0,\xi_0-2\delta_0} \leq & 2B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell+1}\widetilde{\varrho}_0^{-1}. \\ \Big|\frac{\partial\Phi_0}{\partial\boldsymbol{\varphi}}\Big|_{\widetilde{\varrho}_0,\xi_0-2\delta_0} \leq & B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell+1}\delta_0^{-1}. \end{aligned} \tag{5.12.42}$$

Therefore, it makes sense to consider the map

$$\begin{aligned} \mathbf{A} =& \mathbf{A}' + \frac{\partial\Phi_0}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\mathbf{z}), \\ z_j' =& z_j \, \exp\big(i\frac{\partial\Phi_0}{\partial\mathbf{A}'}(\mathbf{A}',\mathbf{z})\big), \qquad j=1,\ldots,\ell \end{aligned} \tag{5.12.43}$$

defined for $(\mathbf{A}',\mathbf{z}) \in C(\widetilde{\varrho}_0, \xi_0-\delta_0; \mathbf{A}_0)$ with values in $\mathcal{C}^{2\ell}$. Here we regard the second of Eqs. (5.12.43) as the complex version of

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}' + \frac{\partial\Phi_0}{\partial\mathbf{A}'}(\mathbf{A}',\mathbf{z}). \tag{5.12.43$'$}$$

Now the problem arises of inverting the first of Eqs. (5.12.43) or the second of Eqs. (5.12.43) in the respective forms

$$\begin{aligned} \mathbf{A}' =& \mathbf{A} + \boldsymbol{\Xi}'(\mathbf{A},\mathbf{z}'), \\ z_j =& z_j' \, \exp\big(i\,\Delta_j(\mathbf{A}',\mathbf{z})\big), \qquad j=(1,\ldots,\ell) \end{aligned} \tag{5.12.44}$$

where the second should be regarded as the complex extension of

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}',\boldsymbol{\varphi}'), \qquad \boldsymbol{\varphi}' \in \mathcal{T}^{\ell}. \tag{5.12.45}$$

For this purpose, we use, respectively, Proposition 21, p.485, and Proposition 20, p.484, §5.11 (choosing, say, $\tau = \log 2$). They guarantee that the above inversions can indeed be made in the desired form, via Eqs. (5.12.39) and (5.12.42), if

$$\frac{1}{2\ell}B_4\varepsilon_0 C\frac{\varrho_0}{\widetilde{\varrho}_0}\delta_0^{-2\ell} \equiv B_4\varepsilon_0 C E_0 C N_0^{\ell+1}\delta_0^{-2\ell} < 1, \tag{5.12.46}$$

[38] Using $\sum_{\boldsymbol{\nu}}|\boldsymbol{\nu}|^a e^{-\delta|\boldsymbol{\nu}|} \leq \max_{y\geq0}(y^a e^{-\delta y/2})\sum_{\boldsymbol{\nu}} e^{-\delta|\boldsymbol{\nu}|/2} \leq \max_{y\geq0}(y^a e^{-y})(\frac{2}{\delta})^a\frac{(4\sqrt{e})^\ell}{\delta^\ell}$ one can take, say, $B_3 = \ell!2^\ell(4\sqrt{e})^\ell$.

where B_4 is a suitable constant determined by imposing Eqs. (5.11.23) and (5.11.27).[39] In this case, $\boldsymbol{\Xi}'$ is holomorphic on $C(\frac{1}{2}\widetilde{\varrho}_0, \xi_0 - 2\delta_0; \mathbf{A}_0)$ as well as $\boldsymbol{\Delta}$ and they verify the bounds

$$\begin{aligned} |\boldsymbol{\Xi}'|_{\widetilde{\varrho}_o/2,\xi_0-2\delta_0} &< B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell}\varepsilon^{2\xi_0} < \frac{\widetilde{\varrho}_0}{8}, \\ |\boldsymbol{\Delta}|_{\widetilde{\varrho}_o/2,\xi_0-2\delta_0} &< 2B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell+1}\widetilde{\varrho}_0^{-1} < \delta_0, \end{aligned} \tag{5.12.47}$$

where the first right-hand-side inequalities follow from Eq. (5.11.29) or Eq. (5.11.25), while the second right-hand-side inequalities follow, if B_4 is chosen as in the footnote,[39] from Eq. (5.12.46).

Eq. (5.12.47) permit us to define on $C(\frac{1}{2}\widetilde{\varrho}_0, \xi_0 - 2\delta_0; \mathbf{A}_0)$, say, the functions

$$\begin{aligned} \boldsymbol{\Xi}(\mathbf{A}', \mathbf{z}') &= \frac{\partial \Phi_0}{\partial \boldsymbol{\varphi}}(\mathbf{A}', \mathbf{z}e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}), \\ \boldsymbol{\Delta}'(\mathbf{A}, \mathbf{z}) &= \frac{\partial \Phi_0}{\partial \mathbf{A}'}(\mathbf{A} + \boldsymbol{\Xi}'(\mathbf{A}, \mathbf{z}), \mathbf{z}) \end{aligned} \tag{5.12.48}$$

and, by Eqs. (5.12.42) and (5.12.39) they verify

$$\begin{aligned} |\boldsymbol{\Xi}|_{\widetilde{\varrho}_o/2,\xi_0-3\delta_0} &< B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell}\varepsilon^{2\xi_0} < \frac{\widetilde{\varrho}_0}{8}, \\ |\boldsymbol{\Delta}'|_{\widetilde{\varrho}_o/4,\xi_0-2\delta_0} &< 4\ell B_3\varepsilon_0 C E_0 C N_0^{\ell+1}\delta_0^{-2\ell+1} < \delta_0. \end{aligned} \tag{5.12.49}$$

Therefore we define the maps $\mathcal{C}_0$, on $(\mathbf{A}', \mathbf{z}') \in C(\frac{1}{2}\widetilde{\varrho}_0, \xi_0 - 2\delta_0; \mathbf{A}_0)$, by

$$\begin{aligned} \mathbf{A} &= \mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}', \mathbf{z}'), \\ \mathbf{z} &= \mathbf{z}'\, e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}. \end{aligned} \tag{5.12.50}$$

and $\widetilde{\mathcal{C}}_0$, on $(\mathbf{A}, \mathbf{z}) \in C(\frac{1}{2}\widetilde{\varrho}_0, \xi_0 - 2\delta_0; \mathbf{A}_0)$, by

$$\begin{aligned} \mathbf{A}' &= \mathbf{A} + \boldsymbol{\Xi}'(\mathbf{A}, \mathbf{z}), \\ \mathbf{z}' &= \mathbf{z}\, e^{i\boldsymbol{\Delta}'(\mathbf{A},\mathbf{z})}. \end{aligned} \tag{5.12.51}$$

which have the properties [by Eqs. (5.12.47) and (5.12.49)]

---

[39] One can take $B_4 = 4\gamma e^2 B_3 \ell$. Note that, not surprisingly, both inversions require the same condition up to a constant factor, adjusted in Eq. (5.12.46) to be the same. This is basically so because the implicit function theorems impose conditions on $\partial^2\Phi_0/\partial\mathbf{A}'\partial\boldsymbol{\varphi}$ for the first inversion or on $\partial^2\Phi_0/\partial\boldsymbol{\varphi}\partial\mathbf{A}'$ for the second.
Actually, it would be easy to check that Eqs. (5.12.42) and (5.12.46) automatically imply that the matrix $J_{ij} = \delta_{ij} + \partial^2\Phi_0/\partial A'_j\varphi_i$ is invertible in $C(\frac{1}{8}\widetilde{\varrho}_0, \xi_0 - 4\delta_0; \mathbf{A}_0)$ if B_3, B_4 are chosen as in footnote 37 to p.498 and as above. Hence, the general theory of the canonical transformations, §3.11, Problems 9-11, p.228, shows that under the condition of Eq. (5.12.46), the map of Eq. (5.12.43) locally generates a completely canonical transformation defined on $\mathcal{S}_{\frac{1}{8}\widetilde{\varrho}_0}(\mathbf{A}_0) \times \mathcal{T}^\ell$, changing $(\mathbf{A}', \boldsymbol{\varphi}')$ into $(\mathbf{A}, \boldsymbol{\varphi})$. This map is actually a globally canonical map, as we shall see.

$$\mathcal{C}_0, \widetilde{\mathcal{C}}_0 : \ C\big(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0\big) \ \to \ C(\frac{1}{2}\widetilde{\varrho}_0, \xi_0 - 2\delta_0; \mathbf{A}_0), \tag{5.12.52}$$

Hence, it makes sense to consider $\mathcal{C}_0\widetilde{\mathcal{C}}_0$ and $\widetilde{\mathcal{C}}_0\mathcal{C}$ on $C(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0)$. By construction, $\mathcal{C}_0$ and $\widetilde{\mathcal{C}}_0$ are inverses of each other:

$$\mathcal{C}_0\widetilde{\mathcal{C}}_0 \equiv \widetilde{\mathcal{C}}_0\mathcal{C}_0 \equiv \{\text{identity map}\} \tag{5.12.53}$$

on $C\big(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0\big)$.

It follows, by the general theory of canonical maps, that $\mathcal{C}_0$ and $\widetilde{\mathcal{C}}_0$ are completely canonical, inverse to each other, maps of $\mathcal{S}_{\frac{1}{4}\widetilde{\varrho}_0} \times \mathcal{T}^\ell$ onto its image.

If a motion takes place in $\mathcal{C}_0(\mathcal{S}_{\frac{1}{4}\widetilde{\varrho}_0} \times \mathcal{T}^\ell)$ it can be described in the $(\mathbf{A}', \boldsymbol{\varphi}')$ variables as a motion generated by the Hamiltonian:

$$H_1(\mathbf{A}', \boldsymbol{\varphi}') = h_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',')) + \mathbf{f_0}(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}','), \boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}', \boldsymbol{\varphi}')) \tag{5.12.54}$$

which, following the perturbation theory and $\forall\,(\mathbf{A}', \mathbf{z}') \in C(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0)$, we write as

$$\begin{aligned} H_1(\mathbf{A}', \boldsymbol{\varphi}') \equiv & \big\{h_0(\mathbf{A}') + f_{0\mathbf{0}}(\mathbf{A}')\big\} \\ & + \big\{h_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}', \boldsymbol{\varphi}')) - h_0(\mathbf{A}') \\ & + f_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}', \boldsymbol{\varphi}'), \boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}', \boldsymbol{\varphi}')) - f_{0\mathbf{0}}(\mathbf{A}')\big\} \\ \stackrel{def}{=} & \{h_1(\mathbf{A}')\} + \{f_1(\mathbf{A}', \boldsymbol{\varphi}')\} \end{aligned} \tag{5.12.55}$$

where h_1 and f_1 are implicitly defined, respectively, as the first and second curly-bracket terms in the intermediate equality in Eq. (5.12.55).

We shall henceforth regard $C(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0)$ as the domain of definition and holomorphy of h_1 and f_1: however we shall further reduce it, later, for the purpose of using dimensional estimates or for other needs. This basic choice of domain is convenient since we control well $\mathcal{C}_0$ on this set, see Eq. (5.12.52).

Our next task, according to the program of the proof, is to find a point $\mathbf{A}_1 \in \mathcal{S}_{\frac{1}{4}\widetilde{\varrho}_0}(\mathbf{A}_0)$ such that

$$\frac{\partial h_1(\mathbf{A}')}{\partial \mathbf{A}'} = \boldsymbol{\omega}_0. \tag{5.12.56}$$

Recalling that $\boldsymbol{\omega}_0 = \frac{\partial h_0(\mathbf{A}_0)}{\partial \mathbf{A}}$ this equation can be elaborated as

$$\begin{aligned}\boldsymbol{\omega}(\mathbf{A}')-\boldsymbol{\omega}(\mathbf{A}_0)+\frac{\partial f_{\mathbf{0}0}}{\partial\mathbf{A}'}(\mathbf{A}')=\mathbf{0}, \qquad &\Rightarrow M(\mathbf{A}-0)(\mathbf{A}'-\mathbf{A}_0)\\ &+\big[\boldsymbol{\omega}(\mathbf{A}')-\boldsymbol{\omega}(\mathbf{A}_0)-M(\mathbf{A}-0)(\mathbf{A}'-\mathbf{A}_0)\\ &+\frac{\partial f_{\mathbf{0}0}}{\partial\mathbf{A}'}(\mathbf{A}')\big]=\mathbf{0}\\ &\Rightarrow(\mathbf{A}'-\mathbf{A}_0)+M(\mathbf{A}_0)^{-1}\big[\boldsymbol{\omega}(\mathbf{A}')-\boldsymbol{\omega}(\mathbf{A}_0)\\ &-M(\mathbf{A}-0)(\mathbf{A}'-\mathbf{A}_0)+\frac{\partial f_{\mathbf{0}0}}{\partial\mathbf{A}'}(\mathbf{A}')\big]=\mathbf{0}\\ &\equiv(\mathbf{A}'-\mathbf{A}_0)+\mathbf{n}(\mathbf{A}')=\mathbf{0}\end{aligned}\tag{5.12.57}$$

where $\mathbf{n}$ is defined on $\widehat{\mathcal{S}}_{\varrho_0}(\mathbf{A}_0)$ by the term within square brackets in the third relation.

Apply Proposition 19, p.484, to the last equation and deduce that if $\gamma|\mathbf{n}|_\varrho<\varrho$ for some $\varrho<\frac{1}{4}\widetilde{\varrho}_0$, then the equation admits a unique solution $\mathbf{A}_1\in\mathcal{S}_\varrho(\mathbf{A}_0)$. Hence, we must estimate $|\mathbf{n}|_\varrho$ for $\varrho<\frac{1}{4}\widetilde{\varrho}_0<\frac{1}{2}\varrho_0$. For this purpose note that

$$\begin{aligned}&|\boldsymbol{\omega}(\mathbf{A}')-\boldsymbol{\omega}(\mathbf{A}_0)-M(\mathbf{A}-0)(\mathbf{A}'-\mathbf{A}_0)|\\ &=\Big|\int_0^1 d\tau\int_0^\tau\delta\theta\frac{d^2}{d\theta^2}\boldsymbol{\omega}(\mathbf{A}_0+\theta\,(\mathbf{A}-\mathbf{A}_0))\Big|\\ &\equiv\Big|\sum_{i,j=1}^{\ell}\int_0^1 d\tau\int_0^\tau\delta\theta\frac{\partial^2\boldsymbol{\omega}}{\partial A_i\partial A_j}(\mathbf{A}_0+\theta\,(\mathbf{A}'-\mathbf{A}_0)(A'_i-A_{0i})(A'_j-A_{0j}))\Big|\\ &\le\varrho^2\ell^2 2\frac{E_0}{(\varrho_0-\varrho)^2}\le 8\ell^2E_0\big(\frac{\varrho}{\varrho_0}\big)^2,\end{aligned}$$

having estimated the second derivative of $\boldsymbol{\omega}$ by a dimensional estimate; see Eqs. (5.11.9) and (5.11.18). Hence, if $\varrho<\frac{1}{4}\widetilde{\varrho}_0$ and if the first of Eqs. (5.12.32) is used with $\boldsymbol{\nu}=\mathbf{0}$:

$$|\mathbf{n}|_\varrho\le\eta_0\left(8\ell^2E_0\big(\frac{\varrho}{\varrho_0}\big)^2+\varepsilon_0\right)\tag{5.12.58}$$

so that if we choose (recalling that $CE_0>l, C\varepsilon_0<1$)

$$\varrho\overset{def}{=}\frac{1}{8\ell}\widetilde{\varrho}_0\sqrt{\frac{\varepsilon_0}{E_0}}<\frac{1}{8}\widetilde{\varrho}_0,$$

it is $|\mathbf{n}|_\varrho<2\varepsilon_0\eta_0$. Applying Proposition 19, p.484, and if $2\gamma\eta_0\varepsilon_0<\frac{1}{8}\widetilde{\varrho}_0$, then Eq. (5.12.56) admits a solution $\mathbf{A}_1\in\mathcal{S}_{\frac{1}{8}\widetilde{\varrho}_0}(\mathbf{A}_0)$. The condition $2\gamma\eta_0\varepsilon_0<\frac{1}{8}\widetilde{\varrho}_0$ becomes, via the expression for $\widetilde{\varrho}_0$ in Eq. (5.12.39), $16\gamma\varepsilon_0C(\eta_0\varrho_0^{-1}E_0)N_0^{\ell+1}<1$; and it can be implied together with Eq. (5.12.46) by requiring

$$B_5\varepsilon_0CE_0C(\eta_0\varrho_0^{-1}E_0)N_0^{\ell+1}\delta_0^{-2\ell}<1,\tag{5.12.59}$$

having used Eq. (5.12.28) and having chosen B_5 suitably,[40] $B_5 > 4$. Hence, if Eq. (5.12.59) holds, the Hamiltonian h, and its perturbation f, in Eq. (5.12.55) can be considered as functions defined and holomorphic in $C(\frac{1}{8}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_1)$ with $\mathbf{A}_1$ so chosen that Eq. (5.12.56) holds.

The argument can now be iterated. In fact it is possible to associate with the Hamiltonians h_1, f_1 in $C(\frac{1}{16}\widetilde{\varrho}_0, \xi_0 - 4\delta_0; \mathbf{A}_1)$[41] the characteristic parameters $\varrho_1 = \frac{1}{16}\widetilde{\varrho}_0, \xi_1 = \xi_0 - 4\delta_0$ and $E_1, \eta_1, \varepsilon_1$, with $E_1, \eta_1, \varepsilon_1$ estimates of

$$\Big|\frac{\partial h_1}{\partial \mathbf{A}'}\Big|_{\varrho_1},\ \Big|\Big|\Big(\frac{\partial^2 h_1}{\partial \mathbf{A}'\partial \mathbf{A}'}\Big)^{-1}\Big|\Big|,\ \Big|\frac{\partial f_1}{\partial \mathbf{A}'}\Big|_{\varrho_1,\xi_1} + \frac{1}{\varrho_1}\Big|\Big|\frac{\partial f_1}{\partial \boldsymbol{\varphi}'}\Big|_{\varrho_1,\xi_1}.$$

To find $E_1, \eta_1, \varepsilon_1$, we apply, as usual, some dimensional estimates. The E_1 estimate is based on the first of Eqs. (5.12.32):

$$\Big|\frac{\partial h_1}{\partial \mathbf{A}'}(\mathbf{A}')\Big| = \Big|\frac{\partial h_0}{\partial \mathbf{A}'}(\mathbf{A}') + \frac{\partial f_{\mathbf{00}}}{\partial \mathbf{A}'}(\mathbf{A}')\Big| \le E_0 + \varepsilon_0. \tag{5.12.60}$$

The η_1 estimate is based on the dimensional estimate, Eq. (5.11.18), for $\sigma_j(\mathbf{A}) = \partial^2 f_1/\partial A'_i \partial A'_j$ as $\sigma_{ij}(\mathbf{A}')| \le \frac{\varepsilon_0}{\varrho_0 - \frac{1}{4}\widetilde{\varrho}_0} \le \frac{2\varepsilon_0}{\varrho_0}$, $\forall\, \mathbf{A}' \in \mathcal{S}_{\varrho_1}(\mathbf{A}_1)$; in fact,

$$\begin{aligned} M_1(\mathbf{A}')^{-1} &= (M_0(\mathbf{A}') + \sigma(\mathbf{A}'))^{-1} = \big(M_0(\mathbf{A}')(1 + M_0(\mathbf{A}')^{-1}\sigma(\mathbf{A}'))\big)^{-1} \\ &= \big(1 + M_0(\mathbf{A}')^{-1}\sigma(\mathbf{A}')\big)^{-1} M_0(\mathbf{A}')^{-1} \\ &\equiv M_0(\mathbf{A}')^{-1}) + \big[\big(1 + M_0(\mathbf{A}')^{-1}\sigma(\mathbf{A}')\big)^{-1} - 1\big] M_0(\mathbf{A}')^{-1}) \end{aligned} \tag{5.12.61}$$

and (since given two $\ell \times \ell$ matrices R and S it is $||RS|| \le ||R||\,||S||$ and $||(1+R)^{-1} - 1|| \le 2||R||$, if $||R|| < \frac{1}{2}$)[42] we see that $||M_0(\mathbf{A}')^{-1}\sigma(\mathbf{A}')|| \le 2\varepsilon_0\eta_0\varrho_0^{-1} < \frac{1}{2}$ [by Eq. (5.12.59)] and

$$||M_1(\mathbf{A}')|| \le \eta_0 + 4\varepsilon_0\eta_0^2\varrho_0^{-1}. \tag{5.12.62}$$

The estimate of ε_1, is slightly more complicated because it involves derivatives of $\boldsymbol{\Xi}$ and $\boldsymbol{\Delta}$: which, however, can be estimated dimensionally. We first elaborate the formal expression of f_1 by adding and subtracting suitable terms:

[40] e.g. one could take $B_5 = 2B_4 16\gamma = 2^{32}\ell!(8\sqrt{e})^\ell \le 2^{32}\ell!!2^{4\ell}$, if $\gamma = 2^8$.

[41] We further restrict the domain in which we consider h_1, f_1 to be able to perform dimensional estimates later.

[42] See **A**ppendix E, Eqs. (E.2) and (E.10), p.523.

$$
\begin{aligned}
f_1(\mathbf{A}',\mathbf{z}) =& h_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}')) - h_0(\mathbf{A}') \\
& + f_0^{[\le N_0]}(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}'), \mathbf{z}' e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}) - f_{0\mathbf{0}}(\mathbf{A}') \\
& + f_0^{[> N_0]}(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}'), \mathbf{z}' e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}) \\
=& \big\{ h_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}')) - h_0(\mathbf{A}') - \boldsymbol{\omega}(\mathbf{A}')\dot{\boldsymbol{\Xi}}(\mathbf{A}',\mathbf{z}') \big\} \\
& + \big\{ f_0^{[\le N_0]}(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}'), \mathbf{z}' e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}) - f_{0\mathbf{0}}(\mathbf{A}') \\
& + \boldsymbol{\omega}(\mathbf{A}')\dot{\boldsymbol{\Xi}}(\mathbf{A}',\mathbf{z}') \big\} \\
& + \big\{ f_0^{[> N_0]}(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}'), \mathbf{z}' e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}) \big\}
\end{aligned}
\tag{5.12.63}
$$

where the addition and subtraction of $\boldsymbol{\omega} \cdot \boldsymbol{\Xi}$ is suggested by the formal perturbation theory and by the fact that, if $(\mathbf{A},\mathbf{z}) = \mathcal{C}_0(\mathbf{A}',\mathbf{z}')$, it is $\boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}') = \frac{\partial \Phi_0}{\partial \boldsymbol{\varphi}}(\mathbf{A}',\mathbf{z})$ so that the various terms in curly brackets formally have size $O(\varepsilon_0^2)$. In fact, the first term is manifestly of $O(|\boldsymbol{\Xi}|^2)$ and $\boldsymbol{\Xi}$ is formally of $O(\varepsilon_0)$; the third term is by construction of formal order $O(\varepsilon_0^2)$, while the second term can be rewritten as

$$
f_0^{[\le N_0]}(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}'), \mathbf{z}' e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')}) - f_0^{[\le N_0]}(\mathbf{A}', \mathbf{z}' e^{i\boldsymbol{\Delta}(\mathbf{A}A',\mathbf{z}')})
\tag{5.12.64}
$$

because, by the definition of Φ_0,

$$
\boldsymbol{\omega}(\mathbf{A}') \cdot \boldsymbol{\Xi}(\mathbf{A}',\mathbf{z}') \equiv \boldsymbol{\omega}(\mathbf{A}') \frac{\partial \dot{\Phi}_0}{\partial \boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi})(\mathbf{A}',\mathbf{z}) = f_0^{[\le N_0]}(\mathbf{A}', \mathbf{z} - f_{0\mathbf{0}}(\mathbf{A}'))
$$

and, therefore, Eq. (5.12.64) is formally of $O(\varepsilon_0)O(\boldsymbol{\Xi})$, i.e., $O(\varepsilon_0^2)$ (here we use that $\mathbf{z}'\, e^{i\boldsymbol{\Delta}(\mathbf{A}',\mathbf{z}')} \equiv \mathbf{z}$, also).

Writing the three terms in curly brackets in Eq. (5.12.63) as $f_0^I, f_0^{II}, f_0^{III}$ respectively, we now show rigorously that they have the right order of magnitude. Dropping the $(\mathbf{A}',\mathbf{z}')$ in the arguments of for simplicity, and using Eq. (5.12.64) and the Taylor-Lagrange formulae, we find

$$
\begin{aligned}
f_1^I =& \int_0^1 dt_1 \int_0^{t_1} dt_2 \, \frac{d^2}{dt_2^2} h_0(\mathbf{A}' + t_2\, \boldsymbol{\Xi}) \\
\equiv& \int_0^1 (1-t)\, dt \Big(\sum_{j,k=1}^{\ell} \frac{\partial^2 h_0}{\partial A_j \partial A_k}(\mathbf{A}' + t\boldsymbol{\Xi}) \Xi_k \Xi_j \Big), \\
f_1^{II} =& \int_0^1 dt\, \frac{d}{dt} f_0^{[\le 0]}(\mathbf{A}' + t\, \boldsymbol{\Xi}, \mathbf{z}' e^{i\boldsymbol{\Delta}}) \\
=& \int_0^1 dt \Big(\sum_{j=1}^{\ell} \frac{\partial f_0^{[\le N_0]}}{\partial A_j}(\mathbf{A}' + t\, \boldsymbol{\Xi}, \mathbf{z}' e^{i\boldsymbol{\Delta}}) \boldsymbol{\Xi}_j \Big), \\
f_1^{III} =& f_o^{[> N_0]}(\mathbf{A}' + t\, \boldsymbol{\Xi}, \mathbf{z}' e^{i\boldsymbol{\Delta}}),
\end{aligned}
\tag{5.12.65}
$$

Bounds for $f_0^I, f_0^{II}, f_0^{III}$ can now be found by dimensional estimates. Combining Eq. (5.11.18) with Eq. (5.12.33), the second of Eqs. (5.12.47) and the first of Eqs. (5.12.49), implies all the following inequalities except the fourth:

$$\begin{aligned}
&\Big|\frac{\partial h_0}{\partial \mathbf{A}\partial \mathbf{A}}\Big|_{\widetilde{\varrho}_0} \le \frac{E_0}{\varrho_o - \widetilde{\varrho}_0} \le 2E_0\varrho_0^{-1},\\
&\Big|\frac{\partial h_0}{\partial \mathbf{A}\partial \mathbf{A}\partial \mathbf{A}}\Big|_{\widetilde{\varrho}_0} \le \frac{2!E_0}{(\varrho_0 - \widetilde{\varrho}_0)^2} \le 6E_0\varrho_0^{-2},\\
&\Big|\frac{\partial f_0^{[\le N_0]}}{\partial \mathbf{A}}\Big|_{\widetilde{\varrho}_0,\xi_0-\delta_0} \le B_1\varepsilon_0\delta_0^{-\ell},\\
&|f_0^{[>N_0]}|_{\widetilde{\varrho}_0,\xi_0-\delta_0} \le \ell B_2\varepsilon_0^2 C\varrho_0,\\
&|\boldsymbol{\Xi}|_{\frac{1}{2}\widetilde{\varrho}_0,\xi_0-3\delta_0} \le B_3\varepsilon_0 C\varrho_0\delta_0^{-2\ell}e^{-2\xi_0} < \frac{\widetilde{\varrho}_o}{8},\\
&|\boldsymbol{\Delta}|_{\frac{1}{2}\widetilde{\varrho}_0,\xi_0-3\delta_0} \le 4\ell B_3\varepsilon_0 CE_0CN_0^{\ell+1}\delta_0^{-2\ell+1} \le \delta_0.
\end{aligned} \tag{5.12.66}$$

To prove the fourth inequality, make use the second of Eqs. (5.12.33):

$$\begin{aligned}
|f_0^{[>N_0]}(\mathbf{A},\mathbf{z})| =&\Big|\sum_{|\boldsymbol{\nu}|>N_0} f_{0\boldsymbol{\nu}}(\mathbf{A})\mathbf{z}^{\boldsymbol{\nu}}\Big| \le \sum_{|\boldsymbol{\nu}|>N_0}|f_{0\boldsymbol{\nu}}|e^{(\xi_0-\delta_0)|\boldsymbol{\nu}|}\\
&\le \sum_{|\boldsymbol{\nu}|>N_0}|\boldsymbol{\nu}|\,|f_{0\boldsymbol{\nu}}|e^{(\xi_0-\delta_0)|\boldsymbol{\nu}|} \le \varepsilon_0\varrho_0\ell\sum_{|\boldsymbol{\nu}|>N_0}e^{-\delta_0|\boldsymbol{\nu}|}\\
&\le\varepsilon_0\varrho_0\ell e^{-\frac{1}{2}N_0\delta_0}\sum_{|\boldsymbol{\nu}|>0}e^{-\frac{1}{2}\delta_0|\boldsymbol{\nu}|} \le B_2\ell\varepsilon_0^2C\varrho_0
\end{aligned} \tag{5.12.67}$$

in $C(\varrho_0, \xi_0 - \delta_0; \mathbf{A}_0)$ [see also (5.12.35)].

For $(\mathbf{A}', \mathbf{z}') \in C(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0)$ it is $(\mathbf{A}' + t\boldsymbol{\Xi}, \mathbf{z}'e^{i\boldsymbol{\Delta}}) \in C(\frac{1}{2}\widetilde{\varrho}_0, \xi_0 - 2\delta_0; \mathbf{A}_0)$; so we can insert the bounds of Eq. (5.12.66) into Eq. (5.12.65), using $e^{\xi} < e < 4$ for simplicity, to obtain

$$\begin{aligned}
|f_1^I|_{\frac{1}{4}\widetilde{\varrho}_o,\xi_0-3\delta_0} \le& 2E_0\varrho_0^{-1}\ell^2(B_3\varepsilon_0C\varrho_0\delta_0^{-2\ell}e^2)\\
&\le 2^9B_3^2\ell^2\varepsilon_0^2CE_0C\varrho_0\delta_0^{-4\ell},\\
|f_1^{II}|_{\frac{1}{4}\widetilde{\varrho}_o,\xi_0-3\delta_0} \le& 2^4B_1B_3\ell\varepsilon_0^2C\delta_0^{-3\ell}\varrho_0,\\
|f_1^{III}|_{\frac{1}{4}\widetilde{\varrho}_o,\xi_0-3\delta_0} \le& B_2\ell\varepsilon_0^2C\varrho_0
\end{aligned} \tag{5.12.68}$$

so that

$$|f_1|_{\frac{1}{4}\widetilde{\varrho}_o,\xi_0-3\delta_0} \le B_6\varepsilon_0^2CE_0C\delta_0^{-4\ell}\varrho_0, \tag{5.12.69}$$

where B_6 is suitably chosen.[43]

[43] e.g. $B_7 = 2^9B_6 = 2^9\ell^2B_3^2 < 2^{4\ell}\ell^2(\ell!)^22^9$.

Next note that $\mathbf{A}_1 \in \mathcal{S}_{\frac{1}{8}\widetilde{\varrho}_0}(\mathbf{A}_0)$ so that $C(\varrho_1, \xi_1; \mathbf{A}_1) \subset C(\frac{3}{16}\widetilde{\varrho}_0, \xi_0 - 4\delta_0; \mathbf{A}_0)$ so that the boundary of $C(\varrho_1, \xi_1; \mathbf{A}_1)$ is quite far from that of $C(\frac{1}{4}\widetilde{\varrho}_0, \xi_0 - 3\delta_0; \mathbf{A}_0)$ and Eq. (5.12.69) yields a dimensional estimate of

$$\widetilde{\varepsilon}_1 \stackrel{def}{=} \sup \Big|\frac{\partial f_1}{\partial \mathbf{A}'}(\mathbf{A}', \mathbf{z}')\Big| + \frac{1}{\varrho_1}\Big|\frac{\partial f_1}{\partial \varphi'}(\mathbf{A}', \mathbf{z}')\Big| \tag{5.12.70}$$

where the supremum is taken over $C(\varrho_1, \xi_1; \mathbf{A}_1)$. It is given by

$$\begin{aligned}
&\widetilde{\varepsilon}_1 \le |f_1|_{\frac{1}{4}\widetilde{\varrho}_o, \xi_0 - 3\delta_0}\Big(\frac{1}{\frac{1}{4}\widetilde{\varrho}_0 - \frac{3}{16}\widetilde{\varrho}_0} + \frac{1}{\varrho_1}\frac{e^{2\delta_0}}{\delta_0}\Big) \\
&\le \frac{2^5 e^2}{\delta_0 \widetilde{\varrho}_0}|f_1|_{\frac{1}{4}\widetilde{\varrho}_o, \xi_0 - 3\delta_0} \le B_7 \varepsilon_0^2 C (E_0 C)^2 N_0^{\ell+1} \delta_0^{-4\ell-1}
\end{aligned} \tag{5.12.71}$$

recalling that $\varrho_1 = \frac{1}{16}\widetilde{\varrho}_0$ and suitably choosing B_7.[44]

So, collecting all the above inequalities (5.12.71), (5.12.62), and (5.12.60) and the definitions of ϱ_1, ξ_1, the following quantities can be taken as characteristic parameters for the Hamiltonians h_1, f_1 in $C(\varrho_1, \xi_1; \mathbf{A}_1)$:

$$\begin{aligned}
\varrho_1 &= \frac{\varrho_0}{32\ell E_0 C N_0^{\ell+1}}, \quad N_0 = \frac{2}{\delta_0}\log\frac{1}{C\varepsilon_0\delta_0^\ell}, \\
\xi_1 &= \xi_0 - 4\delta_0, \\
E_1 &= E_0 + \varepsilon_0, \\
\eta_1 &= \eta_0 + 4\varepsilon_0\eta_0^2\varrho_0^{-1}, \\
\varepsilon_1 &= B_8 C \varepsilon_0^2 (C_0)^2 (\log\frac{1}{C\varepsilon_0\delta_0^\ell})^{\ell+1}\delta_0^{-5\ell-2},
\end{aligned} \tag{5.12.72}$$

having replaced N_0 in Eq. (5.12.71) with its expression in Eq. (5.12.30), and $B_8 = 2^{\ell+1}B_7$ provided the condition of Eq. (5.12.59) holds.

Consider now the mappings $K_n : (\varrho_n, \xi_n, E_n, \eta_n, \varepsilon_n) \to (\varrho_{n+1}, \xi_{n+1}, E_{n+1}, \eta_{n+1}, \varepsilon_{n+1})$ defined by Eq. (5.12.72) in which δ_0 is replaced by δ_n and $0 \to n$, $1 \to n+1$, forgetting (temporarily) the condition of Eq. (5.12.59). Then

$$(\varrho_n, \xi_n, E_n, \eta_n, \varepsilon_n) = K_{n-1}\cdots K_0(\varrho_0, \xi_0, E_0, \eta_0, \varepsilon_0), \tag{5.12.73}$$

and it becomes possible to check that if $C\varepsilon_0$ is small enough (so that the inequality in Eq. (5.12.85) below holds), then:

---

[44] e.g. $B_7 = 2^9 B_6 = 2^{18+4\ell}\ell^2(\ell!)^2$.

$$\begin{aligned}\xi_n >& \xi_\infty = \xi_0 - 4\sum_{j=0}^{\infty}\delta_j\\ E_n \le& 2E_0,\\ \eta_n \le& 2\eta_0,\\ (\varepsilon_0 C)^{(2+\frac{1}{2})^n} \le& \varepsilon_n C \le (\varepsilon_0 C)^{(2-\frac{1}{2})^n},\\ \varrho_n \ge& \frac{\varrho_0}{(E_0C)^n\big[\xi_0^{-n}(n!)^2 2^{\frac{1}{2}n^2}(\log(C\varepsilon_0)^{-1})^{2n}\big]^{\ell+1}}.\end{aligned} \tag{5.12.74}$$

An inductive proof of the validity of Eq. (5.12.74) under a condition of the form of Eq. (5.12.85) is described below, between Eqs. (5.12.75) and (5.12.86), for completeness. The reader should, however, first realize that Eqs. (5.12.74) and (5.12.86) are quite obviously valid under a condition of the type of Eq. (5.12.85) below.

The first inequality follows from our choice of δ_j. The second and third follow from the last two if, say,

$$\varepsilon_0 < \frac{1}{2}E_0, \quad \sum_{n=0}^{\infty}(C\varepsilon_0)^{(\frac{3}{2})^n-1} < 2, \quad \varepsilon_0\eta_0\varrho_0^{-1} < \frac{1}{8}, \tag{5.12.75}$$

$$\sum_{n=0}^{\infty}(C\varepsilon_0)^{(\frac{3}{2})^n-1}\Big[(E_0C)^n[\xi_0^{-n}(n!)^2 2^{\frac{1}{2}n^2}(\log(C\varepsilon_0)^{-1})^{2n}]^{\ell+1}\Big] < 2\log 2$$

The fourth inequality in Eq. (5.12.74) is proved by remarking that, for $x \le 1$, it is $\sup_{0\le x\le 1} x^a(\log\frac{1}{x})^{\ell+1} \le a^{-\ell-1}(\ell+1)!$, $\forall\, a>0$; hence from Eq. (5.12.72),

$$\begin{aligned}\varepsilon_n C\delta_n^\ell \ge& B_8(CE_0)^2(\varepsilon_{n-1}C\delta_{n-1})^2\delta_{n-1}^{-6\ell-2}2^{-\ell},\\ \varepsilon_n C\delta_n^\ell \le& B_8(2CE_0)^2(\varepsilon_{n-1}C\delta_{n-1})^5\delta_{n-1}^{-6\ell-2}3^{\ell+1}(\ell+1)!,\end{aligned} \tag{5.12.76}$$

where the ratio $\frac{\delta_n}{\delta_{n-1}}$ has been bounded by $2^{-\ell}$ or 1 (below or above) and we have applied the above elementary inequality with $a=\frac{1}{3}$.

Since B_8 is very large, e.g., $B_8 2^{-\ell} > 1$, and $CE_0 > 1$, $\delta_{n-1}^{-(6\ell+2)} > 1$ we conclude from the first of Eqs. (5.12.76) that

$$\varepsilon_n C\delta_n^{-\ell} \ge (\varepsilon_{n-1}C\delta_{n-1}^\ell)^2 \ge (\varepsilon_0 C\delta_0^\ell)^{2^n} \tag{5.12.77}$$

which implies the lower bound in Eq. (5.12.74) for $C\varepsilon_n$, if $C\varepsilon_0$ is small enough. And by the explicit expression (5.12.29) for δ_n, the condition turns out to be

$$(C\varepsilon_0)^{-\frac{1}{4}}\delta_0^\ell > 1. \tag{5.12.78}$$

By Eq. (5.12.29) the second inequality in Eqs. (5.12.76) gives

$$\begin{aligned} C\varepsilon_n\delta_n^\ell &\le (C\varepsilon_0\delta_0)^{(\frac{5}{3})^n}\prod_{k=1}^{n} \\ &\cdot\Big[B_8(2CE_0C)^2e^{\ell+1}(\ell+1)!\big(\frac{\xi_0}{16}\big)^{-6\ell-2}(1+n-k)^{12\ell+4}\Big]^{(\frac{5}{3})^{k-1}} \\ &\equiv (C\varepsilon_0\delta_0)^{(\frac{5}{3})^n}\Big[B_8(2CE_0C)^2e^{\ell+1}(\ell+1)!\big(\frac{\xi_0}{16}\big)^{-6\ell-2}\Big]^{((\frac{5}{3})^n-1)\frac{3}{2}} \\ &\cdot e^{(12\ell+4)(\frac{5}{3})^{n-1}\sum_{k=0}^{n-1}(\frac{3}{5})^k\log(1+k)} \\ &\le \Big(C\varepsilon_0\delta_0^\ell(E_0C)^3\xi_0^{-9\ell-3}B_9\Big)^{\left(\frac{5}{3}\right)^n} \end{aligned} \tag{5.12.79}$$

if B_9 is suitably chosen.[45] Since for all $n\ge 0$, $\delta_0^{(\frac{5}{3})^n}\le\delta_n$, Eq. (5.12.79) implies:

$$\begin{aligned} (C\varepsilon_n) &\le (C\varepsilon_0)^{\left(\frac{3}{2}\right)^n}\big[(C\varepsilon_0)^{1-\left(\frac{9}{10}\right)^n}(E_0C)^3\xi_0^{-9\ell-3}B_9\big]^{\left(\frac{5}{3}\right)^n} \\ &\le (C\varepsilon_0)^{\left(\frac{3}{2}\right)^n} \end{aligned} \tag{5.12.80}$$

provided (the worst case being $n=1$)

$$(C\varepsilon_0)^{\frac{1}{10}}(E_0C)^3\xi_0^{-9\ell-3}B_9<1. \tag{5.12.81}$$

Finally consider the last of Eqs. (5.12.74). By the recursive definition of ϱ_n,

$$\varrho_n=\varrho_0\frac{(d_{n-1}\cdots\delta_0)^{\ell+1}}{(2^5E_0C\ell)^{n(\ell+1)}\Big(\prod_{k=1}^n\log\big(C\varepsilon_{n-k}\delta_{n-k}^\ell\big)\Big)^{\ell+1}}, \tag{5.12.82}$$

By Eq. (5.12.77) and the explicit form of δ_n, this becomes

$$\varrho_n\ge\varrho_0\frac{1}{(\ell 2^{5\ell+10}E_0C)^n}\Big[\frac{n!^{-2}\xi_0^n}{2^{\frac{1}{2}n(n-1)}\big(\log(C\varepsilon_0\delta_0^\ell)^{-1}\big)^n}\Big]^{\ell+1}. \tag{5.12.83}$$

So if $C\varepsilon_0$ is small enough, the last inequality in Eqs. (5.12.74) holds. More precisely, it holds if

$$C\varepsilon_0<\delta_0^\ell,\qquad \frac{\big(\log(C\varepsilon_0)^{-1}\big)^{\ell+1}}{2^6\ell+11}\ell>1. \tag{5.12.84}$$

Note that the conditions (5.12.84), (5.12.81), (5.12.78), and (5.12.75) can all be satisfied by imposing a single condition which will also imply Eq. (5.12.59):

[45] e.g. $B_9=8B_83^{\frac{3}{2}(\ell+1)}(\ell+1)!^{\frac{3}{2}}2^6\exp\{\frac{3}{5}(12\ell+4)\sum_{h\ge 0}(\frac{3}{5})^h\log(1+h)\}$.

$$B_{10}\varepsilon_0 C(E_0C)^{\overline{a}}(E_0\eta_0\varrho_0^{-1})^{\overline{b}}\xi_0^{-\overline{c}} < 1, \tag{5.12.85}$$

where B_{10} is a suitable constant and so are $\overline{a}, \overline{b}, \overline{c} > 0$. And if Eq. (5.12.85) holds, then, $\forall\, n \ge 0$, the analogue of Eq. (5.12.59)

$$B_5\varepsilon_n C E_n C(E_n\eta_n\varrho_n)^{-1}N_n^{\ell+1}\delta_n^{-2\ell} < 1 \tag{5.12.86}$$

holds if $C\varepsilon_0$ is small enough. In fact, Eq. (5.12.85) implies Eq. (5.12.74), as just shown, and Eq. (5.12.74) inserted into Eq. (5.12.86) just gives a condition like Eq. (5.12.85) with possibly new values for the constants $B_{10}, \overline{a}, \overline{b}, \overline{c}$.

So a condition of the form of Eq. (5.12.6) guarantees that the sequence of numbers $(\varrho_n, \xi_n, E_n, \eta_n, \varepsilon_n)$ recursively defined in Eq. (5.12.73) verifies Eqs. (5.12.74) and (5.12.86) as well.

This means that under the condition (5.12.6), with B, a, b, c suitably chosen (and ℓ dependent), it is possible to define a sequence of completely canonical transformations, $\mathcal{C}_0, \mathcal{C}_1, \ldots$, having the form

$$\begin{aligned} \mathbf{A} =& \mathbf{A}' + \boldsymbol{\Xi}^{(n)}(\mathbf{A}', \mathbf{z}'), \\ \mathbf{z} =& \mathbf{z}'\, e^{i\boldsymbol{\Delta}^{(n)}(\mathbf{A}', \mathbf{z}')} \end{aligned} \tag{5.12.87}$$

and such that, $\forall\, j = 0, 1, \ldots$,

$$\mathcal{C}_j \,:\, C(\varrho_{j+1}, \xi_{j+1}; \mathbf{A}_{j+1}) \;\to\; C(\varrho_j, \xi_j; \mathbf{A}_j)$$

and [see Eq. (5.12.52) and the discussion following it] Eqs. (5.12.49) and (5.12.47)] $\boldsymbol{\Xi}^{(j)}, \boldsymbol{\Delta}^{(j)}$ can be bounded in $C(\frac{1}{4}\widetilde{\varrho}_j, \xi_j - 3\delta_j; \mathbf{A}_j) \supset C(\varrho_j, \xi_{j+1}, \xi_{j+1}; \mathbf{A}_{j+1})$ by

$$\begin{aligned} &|\mathbf{A}_{j+1} - \mathbf{A}_j| \le \widetilde{\varrho}_j, \\ &|\boldsymbol{\Xi}^{(j)}(\mathbf{A}', \mathbf{z}')| \le B_3\varepsilon_j C\delta_j^{-2\ell}\varrho_j, \\ &|\boldsymbol{\Delta}^{(j)}(\mathbf{A}', \mathbf{z}')| \le 2B_3\varepsilon_j C\delta_j^{-2\ell}\frac{\varrho_j}{\widetilde{\varrho}_j}, \end{aligned} \tag{5.12.88}$$

where $\widetilde{\varrho}_j$ is defined as $\widetilde{\varrho}_0$ with the index j replacing 0 everywhere.

The maps $\mathcal{C}_j$ are very close to the identity map on the very small set on which they are defined. In fact, setting $|(\mathbf{A}, \mathbf{z}) - (\mathbf{A}', \mathbf{z}')| \stackrel{def}{=} |\mathbf{A} - \mathbf{A}'| + \varrho_0|\mathbf{z} - \mathbf{z}'|$, for every pair $(\mathbf{A}', \mathbf{z}'), (\mathbf{A}'', \mathbf{z}'') \in C(\varrho_{j+1}, \xi_{j+1}; \mathbf{A}_{j+1})$ it is:

$$|\mathcal{C}_j(\mathbf{A}', \mathbf{z}') - \mathcal{C}_j(\mathbf{A}'', \mathbf{z}'')| \le (1 + \theta_j)|(\mathbf{A}', \mathbf{z}') - (\mathbf{A}'', \mathbf{z}'')| \tag{5.12.89}$$

where θ_j is a small number that can be taken to be

$$\theta_j = B_{11}\varepsilon_j C\delta_j^{-2\ell-1}\frac{\varrho_j}{\widetilde{\varrho}_j}\frac{\varrho_0}{\widetilde{\varrho}_j} \tag{5.12.90}$$

which is implied by a simple calculation based on the dimensional estimates of the derivatives of $\boldsymbol{\Xi}, \boldsymbol{\Delta}$ on the set $C(\varrho_{j+1}, \xi_{j+1}; \mathbf{A}_j)$ (possible since $\boldsymbol{\Xi}, \boldsymbol{\Delta}$ are

holomorphic on a much larger set, i.e., $C(\frac{1}{4}\widetilde{\varrho}_j,\xi_j-3\delta_j;\mathbf{A}_j))$. Eqs. (5.12.74) imply that $\theta_j \xrightarrow[j\to\infty]{} 0$ very fast, in particular, $\sum_{j=1}^{\infty}\theta_j<\infty$. They also imply that $\sum_{j=1}^{\infty}\theta_j \xrightarrow[\varepsilon_0\to 0]{} 0$. Then a torus can be parametrically defined

$$(\mathbf{A},\mathbf{z})=\mathcal{C}_0\cdots\mathcal{C}_{n-1}\mathcal{C}_n(\mathbf{A}_{n+1},\mathbf{z}'),\qquad \mathbf{z}'\in\mathcal{T}^\ell \tag{5.12.91}$$

which can be written more explicitly as

$$\begin{aligned}\mathbf{A}&=\mathbf{A}_0+\boldsymbol{\alpha}^{(n)}(\boldsymbol{\varphi}')\\ \boldsymbol{\varphi}&=\boldsymbol{\varphi}'+\boldsymbol{\beta}^{(n)}(\boldsymbol{\varphi}'),\end{aligned}\qquad \boldsymbol{\varphi}'\in\mathcal{T}^\ell \tag{5.12.92}$$

where $|\boldsymbol{\Xi}^{(n)}|,|\boldsymbol{\Delta}^{(n)}|$ are defined by comparison between the right-hand sides of Eqs. (5.12.91) and (5.12.92). By construction, $\boldsymbol{\alpha}^{(n)}$ and $\boldsymbol{\beta}^{(n)}$ are holomorphic on the multiannulus $C(\xi_n)\supset C(\xi_\infty)$ and also

$$\begin{aligned}&|\boldsymbol{\alpha}^{(n)}(\mathbf{z}')-\boldsymbol{\alpha}^{(n-1)}(\mathbf{z}')|+\varrho_0|\boldsymbol{\beta}^{(n)}(\mathbf{z}')-\boldsymbol{\beta}^{(n-1)}(\mathbf{z}')|=\\ &=|\mathcal{C}_0\cdots\mathcal{C}_n(\mathbf{A}_{n+1},\mathbf{z}')-\mathcal{C}_0\cdots\mathcal{C}_{n-1}(\mathbf{A}_n,\mathbf{z}')|\\ &\le\Big(\prod_{j=1}^{\infty}(1+\theta_j)\Big)|\mathcal{C}_n(\mathbf{A}_{n+1},\mathbf{z}')-(\mathbf{A}_n,\mathbf{z}')|\\ &\le\Big(\prod_{j=1}^{\infty}(1+\theta_j)\Big)(|\mathbf{A}_{n+1}-\mathbf{A}_n|+|\boldsymbol{\Xi}^{(n)}|+\varrho_0e^{\xi_0}|\boldsymbol{\Delta}^{(n)})\stackrel{def}{=}\sigma_n\end{aligned} \tag{5.12.93}$$

where $|\boldsymbol{\Xi}^{(n)}|,|\boldsymbol{\Delta}^{(n)}|$ denote the right-hand sides of the second and third of Eqs. (5.12.88). Since $\sigma_n \xrightarrow[n\to\infty]{} 0$ very fast, by Eqs. (5.12.88) and (5.12.74), the right-hand side of Eq. (5.12.93) is summable over n. Hence, the limits

$$\boldsymbol{\alpha}_\infty(\boldsymbol{\varphi}')=\lim_{n\to\infty}\boldsymbol{\alpha}^{(n)}(\boldsymbol{\varphi}'),\qquad \boldsymbol{\beta}_\infty(\boldsymbol{\varphi}')=\lim_{n\to\infty}\boldsymbol{\beta}^{(n)}(\boldsymbol{\varphi}') \tag{5.12.94}$$

exist and define (by the convergence theorem of Vitali on the sequences of holomorphic functions) two holomorphic functions of $\boldsymbol{\varphi}'$ in $C(\xi_\infty)$. Via the parametric equations:

$$\begin{aligned}\mathbf{A}&=\mathbf{A}_0+\boldsymbol{\alpha}_\infty(\boldsymbol{\varphi}')\\ \boldsymbol{\varphi}&=\boldsymbol{\varphi}'+\boldsymbol{\beta}_\infty(\boldsymbol{\varphi}'),\end{aligned}\qquad \boldsymbol{\varphi}'\in\mathcal{T}^\ell, \tag{5.12.95}$$

a torus $\mathcal{T}(\boldsymbol{\omega}_0)\subset\mathcal{S}_{\varrho_0}(\mathbf{A}_0)\times\mathcal{T}^\ell$ s defined.

From Eqs. (5.12.88) and (5.12.74) one deduces that $\boldsymbol{\alpha}_\infty$ and $\boldsymbol{\beta}_\infty$ are small if ε_0 is small, i.e., a property like Eq. (5.12.9) holds (possibly redefining B,a,b,c).

So it remains to prove that $\mathcal{T}(\boldsymbol{\omega}_0)$ is an invariant torus run "quasiperiodically" with spectrum $\boldsymbol{\omega}_0$. The Hamiltonian flow $S_t^{(n)}$, which describes in the coordinates defined by the canonical transformation $\mathcal{C}_0\cdots\mathcal{C}_{n-1}$ the perturbed

Hamiltonian flow S_t associated with Eq. (5.12.1), is such that the coordinates of $S_t^{(n)}(\mathbf{A}_n, \boldsymbol{\varphi}')$ of $S_t^{(n)}(\mathbf{A}_n, \boldsymbol{\varphi}')$ are

$$\mathbf{A}_n + \varrho_n O(\varepsilon_n t), \qquad \boldsymbol{\varphi}' + \boldsymbol{\omega}_0 t + O((1 + E_n t)\varepsilon_n t) \tag{5.12.96}$$

because the Hamiltonian f_n contributes terms of order $O(\varepsilon_n)$ to the equations of motion. Of course Eq. (5.12.96) hold only as long as the point in Eq. (5.12.96) is inside $C(\varrho_n, \varepsilon_n; \mathbf{A}_n)$.[46]

If $t > 0$ is fixed, it is clear that $\varrho_n O(\varepsilon_n t) \ll \varrho_n$ for n large, by Eq. (5.12.74), and, therefore, by Eqs. (5.12.89) and (5.12.96), we get

$$\begin{aligned} &|\mathcal{C}_0 \cdots \mathcal{C}_{n-1}(S_t^{(n)}(\mathbf{A}_n, \boldsymbol{\varphi}')) - \mathcal{C}_0 \cdots \mathcal{C}_{n-1}(\mathbf{A}_n, \boldsymbol{\varphi}' + \boldsymbol{\omega}_0 t)| \\ &\quad \le \Big(\prod_{j=1}^{\infty} (1 + \theta_j) \Big) \big(\varrho_n O(\varepsilon_n t) + \varrho_0 O(\varepsilon_n t (1 + E_n t)) \big) \xrightarrow[n \to +\infty]{} 0. \end{aligned} \tag{5.12.97}$$

Hence

$$\begin{aligned} \lim_{n\to\infty} S_t \mathcal{C}_0 \cdots \mathcal{C}_{n-1}(\mathbf{A}_n, \boldsymbol{\varphi}') &\equiv \lim_{n\to\infty} \mathcal{C}_0 \cdots \mathcal{C}_{n-1}(S_t^{(n)}(\mathbf{A}_n, \boldsymbol{\varphi}')) \\ &= \lim_{n\to\infty} \mathcal{C}_0 \cdots \mathcal{C}_{n-1}(\mathbf{A}_n, \boldsymbol{\varphi}' + \boldsymbol{\omega}_0 t) \end{aligned} \tag{5.12.98}$$

but the first and third limit exist by Eqs. (5.12.91) and (5.12.94) and their equality means

$$\begin{aligned} &S_t(\mathbf{A}_0 + \boldsymbol{\alpha}_\infty(\boldsymbol{\varphi}'), \boldsymbol{\varphi}' + \boldsymbol{\beta}_\infty(\boldsymbol{\varphi}')) \\ &\quad = (\mathbf{A}_0 + \boldsymbol{\alpha}_\infty(\boldsymbol{\varphi}' + \boldsymbol{\omega}_0\, t), \boldsymbol{\varphi}' + \boldsymbol{\omega}_0\, t + Bb_\infty(Bf' + \boldsymbol{\omega}_0\, t)) \end{aligned} \tag{5.12.99}$$

which just says that $\mathcal{T}(\boldsymbol{\omega}_0)$ is an invariant torus for the perturbed motion on which quasi-periodic motions with spectrum $\boldsymbol{\omega}_0$ take place ($t > 0$ being arbitrary). mbe

5.12.1 Problems

1. Let $\mathbf{A} \in \mathcal{S}_\varrho(\mathbf{A}_0)$ and write Eq. (5.12.4) for $\boldsymbol{\nu} = \mathbf{e}^{(1)} = (1, 0, \ldots, 0)$ as $|\boldsymbol{\omega}_1(\mathbf{A})|^{-1} \le C$ Deduce that this implies $E_0 C > 1$, with the assumptions and notations of, say, Proposition

[46] One finds this as follows: let $S_t^{(n)}(VA_n, \boldsymbol{\varphi}') = (\mathbf{A}(t), \boldsymbol{\varphi}(t))$ so that

$$\dot{\mathbf{A}} = -\frac{\partial f_n}{\partial \boldsymbol{\varphi}}, \dot{\boldsymbol{\varphi}}' = \boldsymbol{\omega}_n(\mathbf{A}) + \frac{\partial f_n}{\partial \mathbf{A}} \equiv \boldsymbol{\omega}_0 + (\boldsymbol{\omega}_n(\mathbf{A}) - \boldsymbol{\omega}_0) + \frac{\partial f_n}{\partial \mathbf{A}}$$

Therefore by Taylor's theorem and by a dimensional estimate, one finds after integration over t

$$|\mathbf{A}(t) - \mathbf{A}_n| \le \varepsilon_n \varrho_n t, \qquad |\boldsymbol{\varphi}'(t) - \boldsymbol{\varphi}' - \boldsymbol{\omega}_0 t| \le (E_n \widetilde{\varrho}_n^{-1} \varrho_n \varepsilon_n t)\, t + \varepsilon_n t$$

which implies Eq. (5.12.96).

22. In the above context show that $M_{ij}(\mathbf{A}_0)| \le \frac{E_0}{\varrho_0}$, by a dimensional estimate [see Eq. (5.11.18)], and deduce from this that $\ell |M_{\varrho_0}^{-1} E_0 > \varrho_0$. (*Hint:* $1 \equiv (|M(\mathbf{A})^{-1} M(\mathbf{A}))_{11} = |\sum_{k=1}^{\ell} |M(\mathbf{A})_{1k}^{-1} M(\mathbf{A})_{k1}| \le \frac{E-0}{\varrho_0} \ell |M(\mathbf{A})^{-1}|$ or $\le \frac{E_0}{\varrho_0} ||M(\mathbf{A})^{-1}|| \ldots$.)

2. Consider the Hamiltonian on $\mathcal{R}^{d+1} \times \mathcal{R}^{d+1}$:

$$\frac{A^2}{2} + \frac{\mathbf{B}^2}{2} + \varepsilon f(\varphi, \boldsymbol{\psi}), \qquad (\varphi, \boldsymbol{\psi}) \in \mathcal{T}^1 \times \mathcal{T}^d, \quad (A, \mathbf{B}) \in \mathcal{R}^1 \times \mathcal{R}^d$$

Consider the motions near the resonating torus $A = 1, \mathbf{B} = \mathbf{0}$ and write

$$A = 1 + \sqrt{\varepsilon}\, a_\varepsilon(t\sqrt{\varepsilon}), \qquad \varphi = \delta_\varepsilon(t\sqrt{\varepsilon}),$$
$$\mathbf{B} = \sqrt{\varepsilon}\, \mathbf{b}_\varepsilon(t\sqrt{\varepsilon}), \qquad \boldsymbol{\psi} = \boldsymbol{\gamma}_\varepsilon(t\sqrt{\varepsilon})$$

for the solution to the Hamiltonian equations with initial datum

$$a_\varepsilon(0) = a_0, \ \mathbf{b}_\varepsilon(0) = \mathbf{b}_0, \ \boldsymbol{\gamma}_\varepsilon(0) = \boldsymbol{\gamma}_0, \ \delta_\varepsilon(0) = \delta_0.$$

Show that the solutions to the Hamiltonian equations are such that $a_\varepsilon, \mathbf{b}_\varepsilon, \boldsymbol{\gamma}_\varepsilon$ (but not δ_ε) have a limit as $\varepsilon \to 0$ and this limit verifies the equations

$$\dot{a} = 0, \ \dot{\boldsymbol{\gamma}} = \mathbf{b}, \ \dot{\mathbf{b}} = -\frac{\partial \overline{f}(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}},$$

where $\overline{f}(\boldsymbol{\gamma}) = \int_0^{2\pi} f(\theta, \boldsymbol{\gamma}) \frac{d\theta}{2\pi}$. Show that the limit is approached with a speed $O(\varepsilon t)$ at fixed t. (*Hint:* Write the Hamiltonian equations and note that, after dividing them by $\sqrt{\varepsilon}$, they converge formally to the above equations for $a, \mathbf{b}, \gamma$. Then apply the ideas of the proof of Proposition 13, p.186, §3.8, and of §3.7 and §3.8.)

3. In the context of Problem 2, take $d = 1$. Show that "up to a time $O(1/\varepsilon)$", the motion is quasi-periodic with pulsations $\omega_1 = 1, \omega_2 = \frac{2\pi}{T_{b_0,\gamma_0}} \sqrt{\varepsilon}$, where

$$T_{\mathbf{b}_0, \boldsymbol{\gamma}_0} = 2 \int_{\gamma_-}^{\gamma_+} \frac{d\gamma}{\sqrt{2(E - \overline{f}(\gamma))}}$$

where $E_0 = \frac{1}{2} b_0^2 + \overline{f}(\gamma_0)$ and γ_-, γ_+ are 0 and 2π if the equation $E_0 = \overline{f}(\gamma)$ has no roots; otherwise, they are two suitably chosen roots of this equation. Consider only the case $T_{b_0,\gamma_0} < +\infty$ (however, the data for which $T_{b_0,\gamma_0} = +\infty$ are exceptional).

4. Find a result analogous to the one of Problem 2 near a general torus with rational pulsations for the solution flow of the equations associated with the Hamiltonian

$$\frac{1}{2}\mathbf{A}^2 + \varepsilon\, f(\boldsymbol{\varphi}).$$

(*Hint:* First extend Problem 2 to the case when f depends on $\mathbf{A}, \mathbf{B}$ also; then canonically change variables so that the torus under analysis appears to be run with pulsations $\boldsymbol{\omega} = (\omega_0, 0, 0, \ldots, 0)$.)

5. Consider a time-dependent Hamiltonian with one degree of freedom, periodic in time with period 2π: $h_0(\mathbf{A}) + f_0(\mathbf{A}, \boldsymbol{\varphi}, t)$; see Problems 12-14, p.478, §5.10.
Suppose that h_0 is holomorphic in $\widehat{\mathcal{S}}_{\varrho_0}(A_0)$ and that f_0 is holomorphic in $C(\varrho_0, \xi_0, A_0) = \widehat{\mathcal{S}}_{\varrho_0}(\mathbf{A}_0) \times C(\xi_0)^2 = \{A, \boldsymbol{\varphi}, t | (A, z, \zeta) \in \mathcal{C}^3, \, |A - A_0| < \varrho, e^{-\xi_0} < |z| < e^{\xi_0}, e^{-\xi_0} < |\zeta| < e^{\xi_0}$, where $z = e^{i\varphi}, \zeta = e^{it}$. Using the formal perturbation theory of Problem 12, §5.10, p.478, prove that if $\frac{dh_0}{dA} \ne 0$ and f_0 is "small" and

$$|\omega(A_0)\nu_1 + \nu_2|^{-1} \le C\,(|\nu_1| + |\nu_2|)^\alpha, \qquad \forall\, \mathbf{0} \ne \boldsymbol{\nu} \in \mathcal{Z}^2,$$

for some $C, \alpha > 0$, then the perturbed motion, regarded as taking place on the space of the variables $(\mathbf{A}, \varphi, t)$, leaves invariant a torus on which a quasi-periodic motion with pulsations $(\omega(A_0), 1)$ takes place in the following sense. There exist two holomorphic functions $\alpha_\infty, \beta_\infty$ on $\mathcal{T}^2$ such that setting $A = A_0 + \alpha_\infty(\varphi', t')$, $\varphi = \varphi' + \beta_\infty(\varphi', t')$, $t = t'$, the solution of the equations of motion with datum assigned at time t' and given by $A = A_0 + \alpha_\infty(\varphi', t')$, $\varphi = \varphi' + \beta_\infty(\varphi', t')$ for some $\varphi' \in \mathcal{T}^1$ evolves at time $t' + \tau$ into

$$\begin{aligned} A(\tau) =& A_0 + \alpha_\infty(\varphi' + \omega\tau, t' + \tau), \\ \varphi(\tau) =& \varphi' + \omega\tau + \beta_\infty(\varphi' + \omega\tau, t' + \tau), \end{aligned}$$

i.e., regarding the phase space as $\mathcal{R} \times \mathcal{T}^2$, the above motions can be regarded as taking place on a two-dimensional torus in $\mathcal{R} \times \mathcal{T}^2$ and having pulsations $(\omega, 1)$. (*Hint:* Just repeat the proof of Proposition 22. No real simplification arises in this apparently simpler case.)

6. Consider the Hamiltonian ("Duffing oscillator") $H = \frac{1}{2}p^2 + \frac{1}{4}q^4 + \varepsilon\, q \sin t$. Fix an initial datum (p_0, q_0). Show that if ε is small enough, the trajectory with datum (p_0, q_0) at any initial time t_0 is uniformly bounded in time. (*Hint:* Show that p_0, q_0 is between two unperturbed tori in the phase space $\mathcal{R}^1 \times \mathcal{T}^2$, of the system with $\varepsilon = 0$, having pulsations $(\omega_1, 1), (\omega_2, 1)$ (see preceding problem) nonresonant and with finite resonance parameter C. Use Problem 5 to show that for ε small, such tori are slightly deformed but remain invariant. Then use the fact that a two-dimensional torus in a three-dimensional space has an "interior" and an "exterior".)

7. In the context of Problem 5, define $\boldsymbol{\varphi} = (\varphi, t)$ and $E_0 \ge \left|\frac{dh}{dA}\right|_{\varrho_0}$, $\eta_0 \ge \left|\left(\frac{d^2h}{dA^2}\right)^{-1}\right|_{\varrho_0}$, $\varepsilon_0 \ge \left|\frac{\partial f}{\partial A}|_{\varrho_0,\xi_0} + \frac{1}{\varrho_0}\right|\frac{\partial f}{\partial \varphi}|_{\varrho_0,\xi_0}$. Then the condition of smallness of ε_0 for the property envisioned there is implied by the following condition, as can be proven:[47]

$$10^{20}(\eta_0 E_0 \varrho_0^{-1})^4 (C E_0)^4 C \varepsilon_0 < 1.$$

Derive a similar formula (i.e., prove the statement in Problem 5, explicitly computing the constants) and try to improve it.

8. Consider the system on $\mathcal{R} \times \mathcal{T}^2$ ("Escande-Doveil pendulum")

$$\frac{1}{2}A^2 + \varepsilon\,(\cos\varphi + \cos(\varphi - t)),$$

where t is the time; see Problems 12-14, §5.10, p.478. Apply the result of Problem 5 with the estimate in Problem 7 to place a bound on how large ε must be in order that one cannot guarantee the "stability of the quasi-periodic motions with pulsations $(\omega_0, 1)$" with $\omega_0 = \frac{1}{2}(\sqrt{5} - 1) = \{\text{golden section}\}$.

9. Same as Problem 8, but applying the results of Problems 5 and 7 to the system obtained from the one in Problem 8, by first removing the perturbation to $O(\varepsilon)$ by ordinary perturbation theory; see Problems 12-14, §5.10, p.478.

10. Same as Problem 9, but first removing the perturbation to $O(\varepsilon^4)$. Check that in this way one obtains much better results.

11. Suppose that, observing the motions of the system in Problem 8, one is able to see them with an absolute precision η of four digits (in decimal basis) and for an observation time T about equal to 50 periods of the forcing term, $T = 50 \cdot 2\pi$.
Note that (see (5.12.86)) to achieve a given accuracy for a given time, one only needs to

[47] [23].

"remove the perturbation" to an order n such that $O(\varepsilon_n T(1+E_n T)) < \eta$. Using this remark, estimate a threshold for the "survival" of motions which look quasi-periodic, within the error η up to time T, with pulsations $(\omega_0, 1)$, $\omega_0 = \frac{1}{2}(\sqrt{5}-1)$, and compare the result with the experimental value of the "threshold of disappearance" of the quasi-periodic motion in question: $\varepsilon \simeq 0.75$.

12. Try to compute the constants B, a, b, c in Eq. (5.12.6), explicitly improving the values of the constants B_1-B_9 suggested in the proof of Proposition 22. An example of a rigorous result is[48]

$$\ell^{12\ell} 10^{40\ell} (\eta_0 E_0 \varrho_0^{-1} C E_0)^{14} \xi_0^{-2(10\ell+6)} < 1.$$

The following problems constitute a follow up of the problems in §4.10 on the theory of precession. None of the approximations suggested below for performing the lowest order perturbation theory is, strictly speaking necessary: the calculations could be easily carried out without any approximation at all, leading essentially to the same results. They would however be extremely cumbersome. In practice they have never been done because already with the approximations below it is clear that one has reached a precision where the non rigid structure of the Earth is important togheter with its density irregularities, and the consequent non rotationally symmetric shape: therefore the use made of the following calculations is just to provide some formulae with free parameters to be used to perform numerical fits in the tables, much in the same spirit that animated the Greek astronomy (which is not a good reason for not trying someday a better calculation to test if newtonian mechanics can be applied to the theory of nutation to investigate the elastic properties of the planet). Let $\bar{\omega}_D, \bar{\imath}$ be some approximations of the mean daily rotation angular velocity and of the mean inclination of Earth's axis. Below we suppose, as it is the case for the Earth, that $1 >> (1-L^2/A^2)^{1/2} >> \bar{\omega}_p/\bar{\omega}_D$ (where we call $\bar{\omega}_p$ the precession velocity calculated from the formula of problem (16) of §4.10, with such approximate values for the Earth angular velocity and inclination); this means that for many purposes the axis of rotation, the axis of symmetry and the axis of the angular momentum of the Earth can be confused, even though the theory is precisely looking for phenomena that exist just because such axes are not identical.

13. In the context of problems (8) through (7) of §4.10, show that the Hamiltonian in §4.10, problem (14), can be written, if $K \equiv K_z$:

$$\begin{aligned} H_p = & \frac{A^2}{2J} + \frac{3\eta_1 k M_S}{2a^3} J\Big((1-\frac{K^2}{A^2})\sin^2(\lambda_T-\gamma) + \Big(1-\frac{K^2}{A^2}\Big)^{1/2}\Big(1-\frac{L^2}{A^2}\Big)^{1/2} \\ & \cdot\Big((\frac{K}{A}-1)\sin(\lambda_T-\gamma+\varphi) + (\frac{K}{A}+1)\sin(\lambda_T-\gamma-\varphi)\Big)\Big) \end{aligned}$$

when one neglects, in the perturbation terms, $(1-L^2/A^2) \equiv \sin^2\theta$ (but not its square root) and the eccentricity of the Earth orbit. This means, as it appears below, that we neglect the difference between the Earth angular momentum axis, the Earth instantaneous rotation

[48] see [11].

direction and the Earth symmetry axis everywhere in the hamiltonian except in the places where they will produce the largest corrections to the equations of motion.

14. Show that neglecting terms proportional to $(1 - L^2/A^2)$ as well as the variability of the Earth axis the hamiltonian in problem (13) can be put, using the notations of the problems of §4.10, in the form:

$$\begin{aligned}H_p =&\omega_T B + \frac{A^2}{2J} - \frac{3}{2}\eta_1\omega_T^2\frac{JK^2}{2A^2} + \frac{3}{2}\eta_1\omega_T^2 J\Big\{-\frac{1}{2}\Big(1 - \frac{K^2}{A^2}\Big)\cos 2(\lambda-\gamma) + \\ &+ [(\frac{K}{A} - 1)\sin(\lambda - \gamma + \varphi) + (\frac{K}{A} + 1)\sin(\lambda - \gamma - \varphi)]\Big(1 - \frac{K^2}{A^2}\Big)^{1/2}\Big(1 - \frac{L^2}{A^2}\Big)^{1/2}\Big\}\end{aligned}$$

Note that L is a constant of motion and therefore it will not be considered a canonical variable; the new canonical coordinates (B, λ) have been introduced artificially to make the system autonomous. Note also that the parameters $\bar{A}, \bar{\imath}$ are fictitious parameters, so far, as they drop out of the above formula if $\bar{\omega}_p$ is fully rexpressed in terms the constants in problem (13). (*Hint:* note that if the orbit is regarded as circular then λ_T can be identified up to an additive constant with the average anomaly; hence it rotates at constant rate ω_T; the auxiliary variable B will play no role here).

15. The classical theory of nutation averages the Hamiltonian in problem (14) over the fast angles φ, but *not* over the relatively slower angles λ or over the very slow γ. The Hamiltonian thus obtained should reliably describe motions over a time scale $\gg 2\pi/\omega_D = 1$ day and it is:

$$H_D = \frac{A^2}{I_3} + \frac{3}{2}\eta_1\omega_T^2 J\left(1 - \frac{K^2}{A^2}\right)\sin^2(\lambda - \gamma)$$

Show that this is integrable by quadratures and, setting $\gamma - \omega_T t = \tilde{\gamma}$, reducible to the quadrature:

$$\begin{aligned}&\int_{\tilde{\gamma}_0}^{\tilde{\gamma}} \frac{d\gamma'}{-\omega_T - \frac{3}{2}I_3\eta_1\omega_T^2(2K/A^2)\sin^2\tilde{\gamma}'} = t - t_0\\ &\frac{3}{2}\eta_2 I_3\omega_T^2(1 - \frac{K^2}{A^2})\sin^2\tilde{\gamma} - \omega_T K = -\omega_T K_0\end{aligned}$$

having called $K_0, \tilde{\gamma}_0, t_0$ the values of $K, \tilde{\gamma}, t$ when $\gamma - \omega_T t = 0$. Show that, neglecting the variations of K of higher order in η one finds that the motion is:

$$\begin{aligned}\dot{\gamma} =& -\omega_p t - \frac{3}{2}\omega_T^2 J\eta_1\frac{2K_0}{A^2}(\sin^2(\omega_T t + \lambda_0 - \gamma_0) - \frac{1}{2})\\ \dot{K} =& -\frac{3}{2}\omega_T^2 J\eta_1(1 - \frac{K_0^2}{A^2})\sin 2(\omega_T t + \lambda_0 - \gamma_0)\end{aligned}$$

hence, recalling that $\cos\delta = K/a$ and writing $\delta = i_0 + \delta'$ and $\gamma + \omega_p t = \gamma'$, it is:

$$\delta' = \frac{3}{2}\eta\left(\frac{\omega_T}{\omega_D}\right)\sin\delta_0\cos 2\omega_T, \qquad \gamma' = \frac{3}{2}\eta\left(\frac{\omega_T}{\omega_D}\right)\cos\delta_0\cos 2\omega_T$$

and we see that the two Euler angles expressing the deviations from the *mean* precession motion move on a small ellipse with a period equal, in this approximation, to $2\pi/\omega_T$. This is the solar *nutation motion.*

16. If the Moon is taken into account along a similar scheme one finds that the Moon nutation makes the δ', γ' revolve still over an ellipse. The theory has to be done from the beginning as the main cause of the nutation due to the Moon is the fact that the plane of motion of the Moon is not fixed in space but has a precession on a cone of angle equal to the moon inclination $i_L \sim 5^o$ with a period of the $2\pi/\omega_{pL}$, for some ω_L. The nutation due to the

Moon comes out to be on an ellipse about 10 times larger than that calculated above for the Sun contribution and has a period of the order of $2\pi/\omega_{pL}$. Check that such period has the order of 20 years. (*Hint:* The precession of the Moon plane is due to the gravitational force of the Sun. One can imagine, for the purpose of studying phenomena that take place over a time scale large with respect to the Moon period of revolution ($T_L =\sim 27$ days) that the Moon is uniformly spread on its orbit on an annulus of radius a_L whose plane is inclined of i_L to the ecliptic and which is rotating around its center T at velocity ω_L equal to the mean angular velocity of the Moon $\omega_L = 2\pi/T_L$. The annulus is at a distance a from the Sun and gravitates around it with angular velocity ω_T, (neglecting the eccentricities of Earth and Moon), hence it has a precession that can be calculated from that of the Earth simply by using the value η_1 appropriate for an annulus, *i.e.* 1/2 as the inertia moments of an annulus are $J = M_L a_L^2$ and $I = J/2$. Hence the precession velocity is $\omega_{pL} = -(3/4)\omega_T^2\omega_L^{-1}$.)

18. Show that the generating function of the canonical map formally removing, from the above H_p, the perturbation to higher order is $A_0\varphi + K_0\gamma + B_0\lambda + \Phi$ with:

$$\begin{aligned}\Phi(A_0, K_0, \varphi, \gamma, \lambda) = &-\frac{\bar\omega_p \bar A}{\cos\bar\imath}\Big\{-\frac{1}{4\omega_T}(1-\frac{K_0^2}{A_0^2})\frac{\sin 2(\lambda-\gamma)}{1-\omega_p/\omega_T} - \frac{1}{\omega_D}\\ &\cdot\Big(1-\frac{K_0^2}{A_0^2}\Big)^{1/2}\Big(1-\frac{L^2}{A_0^2}\Big)^{1/2}\cdot\Big[(\frac{K_0}{A_0}-1)\frac{\cos(\lambda-\gamma+\varphi)}{1+(\omega_T-\omega_p)/\omega_D}\\ &-(\frac{K_0}{A_0}+1)\frac{\cos(\lambda-\gamma-\varphi)}{1-(\omega_T-\omega_p)/\omega_D}\Big]\Big\}\end{aligned}$$

where $\omega_D \equiv A_0/I_3$, and $\omega_p \equiv \bar\omega_p \bar A/2A_0^2\cos\bar\imath$.

19. Consider the canonical map with generating function $A_0\varphi+K_0\gamma+B_0\lambda+\Phi(A_0, K_0, \varphi, \gamma)$ and introduce the parameters $\varepsilon = (1 - L^2/A_0{}^2)^{1/2}$, $\cos i_0 \equiv K_0/A_0$ and set $q_\pm = (1 \pm \cos i_0)/\cos i_0$. To simplify the calculations choose the so far arbitray constants $\bar\imath, \bar\omega_D$ to be identical to $i_0, \omega_D \equiv A_0/I_3$. Show that with this choice of $\bar\imath, \bar\omega_D$ the map is generated by the relations:

$$\begin{aligned}\varphi_0 =&\varphi - \frac{\omega_p}{\omega_D\varepsilon}\Big(q_+\frac{\cos(\lambda-\gamma-\varphi)}{1-(\omega_T-\omega_p)/\omega_D} + q_-\frac{\cos(\lambda-\gamma+\varphi)}{1+(\omega_T-\omega_p)/\omega_D}\Big)\\ \gamma_0 =&\gamma - \frac{\omega_p}{\omega_T}\frac{\sin 2(\lambda-\gamma)}{2(1-\omega_p/\omega_T)}\\ A =&A_0 - \frac{\omega_p A_0}{\omega_D}\varepsilon\sin i_0\Big(q_+\frac{\sin(\lambda-\gamma-\varphi)}{1-(\omega_T-\omega_p)/\omega_D} - q_-\frac{\sin 2(\lambda-\gamma+\varphi)}{1+(\omega_T-\omega_p)/\omega_D}\Big)\\ K =&K_0 - \frac{\omega_p}{2\omega_T}A_0\tan i_0\frac{\cos 2(\lambda-\gamma)}{1-\omega_p/\omega_T}\end{aligned}$$

and, trivially, $\lambda_0 \equiv \lambda$. Neglecting terms of order $O((\omega_p/\omega_T)^2)$ as well as terms of order $O(\omega_p\varepsilon/\omega_D)$ and assuming that $\omega_p/\varepsilon\omega_D\tan i_0$ is very small the above relations are trivially inverted, up to corrections of higher order, as:

$$\begin{aligned}\varphi =&\varphi_0 + \frac{\omega_p}{\omega_D\varepsilon}\Big(q_+\frac{\cos(\lambda_0-\gamma_0-\varphi_0)}{1-(\omega_T-\omega_p)/\omega_D} + q_-\frac{\cos(\lambda_0-\gamma_0+\varphi_0)}{(1+(\omega_T-\omega_p)/\omega_D)}\Big)\\ \gamma =&\gamma_0 - \frac{\omega_p}{2\omega_T}\frac{\sin 2(\lambda_0-\gamma_0)}{(1-\omega_p/\omega_T)}\\ A =&A_0\Big(1 - \frac{\omega_p}{4\omega_D}\varepsilon\sin i_0\Big(q_+\frac{\sin(\lambda_0-\gamma_0-\varphi_0)}{1-(\omega_T-\omega_p)/\omega_D} - q_-\frac{\sin(\lambda_0-\gamma_0+\varphi_0)}{1+(\omega_T-\omega_p)/\omega_D}\Big)\Big)\\ K =&K_0\Big(1 + \frac{\omega_p}{2\omega_T\cos i_0}\tan i_0\frac{\cos 2(\lambda_0-\gamma_0)}{1-\omega_p/\omega_T}\Big)\end{aligned}$$

and the equations of motion are, in the new coordinates labeled by 0:

$$\dot\varphi_0 = \omega_D \qquad \dot\gamma_0 = \omega_p$$
$$A_0 = I_3\omega_D \qquad K_0 = I_3\omega_D \cos i_0$$

and $i_0, \omega_D, \omega_p, \varepsilon$, as well as the initial data for $\varphi_0, \gamma_0, \lambda_0$, must be regarded as parameters to be determined from observations. They define the *mean inclination, daily rotation, equinox precession, and nutation constant.*

Show that the terms neglected in problems (13) thorugh (15) and above would add oscillating terms with much smaller amplitude.

20. Consider the motions described in the new primed coordinates by the last equations of problem (18). The angles φ, γ can be thought to be animated by two distinct motions. The first are the two *precession* motions:

$$\varphi \to \varphi + \omega_D t, \qquad \gamma \to \gamma + \omega_p t$$

are, respectively, the *mean daily rotation* and the *mean precession of the equinoxes*. The second motion is linearly superposed to the first and is the motion obtained by replacing λ by $\lambda + \omega_T t$, φ_0 by $\varphi_0 + \omega_D t$ and γ_0 by $\gamma_0 + \omega_p t$ in the trigonometric terms in the second of problem (18). The second motion is the *nutation* caused by the Sun.

21. The axis $\mathbf{k}_0$ with Euler angles $(i_0, \gamma_0 + \omega_p t)$ is called the *mean axis of rotation*: it is an axis animated by a purely precessional, uniform, motion. Its node $\mathbf{m}_0$ on the ecliptic plane is the *mean node* or *mean equinox* and it is rotating at uniform angular velocity ω_p. Show that, in the approximation in which the Earth axis, the angular momentum axis and the angular velocity axis are identified, the actual inclination i of the axis and the longitude δ of the apparent (*i.e.* the actual) node with respect to the mean node are given by:

$$\begin{aligned}
\cos i =& \frac{K}{A} = \cos i_0 \Big(1 + \frac{\omega_p}{2\omega_T \cos i_0} \tan i_0 \, \frac{\cos 2(\lambda - \gamma_0)}{1 - \omega_p/\omega_T} + \\
&+ \frac{\omega_p}{2\omega_D} \varepsilon \sin i_0 (q_- \frac{\sin(\lambda - \gamma_0 - \varphi_0)}{1 - (\omega_T - \omega_p)/\omega_D}) - q_+ \frac{\sin(\lambda - \gamma_0 + \varphi_0)}{1 + (\omega_T - \omega_p)/\omega_D}\Big) \\
\delta =& \frac{\omega_p}{2\omega_T} \tan i_0 \, \frac{\sin 2(\lambda - \gamma_0)}{1 - \omega_p/\omega_T}
\end{aligned}$$

having set $q_\pm = (1 \pm \cos i_0)/\cos i_0$.

22. Consider the mean Earth axis and a plane orthogonal to it. Show that the coordinates of $\mathbf{i}_3$, the actual rotation axis, and those of the mean rotation axis and mean equinox, $\mathbf{k}_0, \mathbf{m}_0$ are:

$$\begin{aligned}
\mathbf{i}_3 =& (\sin i \sin(\gamma_0 + \delta), -\sin i \cos(\gamma_0 + \delta), \cos i) \\
\mathbf{m}_0 =& (\cos \gamma_0, \sin \gamma_0, 0) \\
\mathbf{k}_0 =& (\sin i_0 \sin \gamma_0, \sin i_0 \cos \gamma_0, \cos i_0)
\end{aligned}$$

Show that the coordinates of the extreme point projected on the just constructed plane are given by:

$$\begin{aligned}
x =& \mathbf{i}_3 \cdot \mathbf{m}_0 = \sin \delta \sin i \\
y =& \mathbf{i}_3 \cdot \mathbf{k}_0 \wedge \mathbf{m}_0 = \sin \gamma_0 (\cos i \sin i_0 \sin \gamma_0 - \cos i_0 \sin i \sin(\gamma_0 + \delta)) - \\
& - \cos \gamma_0 (\cos i_0 \sin i \cos(\gamma_0 + \delta) - \cos i \sin i_0 \cos i \cos \gamma_0)
\end{aligned}$$

Show that if in the equations of motion one neglects the terms with the angles $(\lambda - \gamma_0 \pm \varphi_0)$ then the endpoint of the rotation axis describes an ellipse, *i.e.* (x, y) describe an ellipse.

23. In general one has to take into account the force of the Moon, which in fact produces terms greater than the ones considered above from the Sun. However the motions will be basically of the same type: the nutation and precession will receive contributions also from the Moon and the other planets. If one really wants, one can improve the above description by distinguishing the three main rotation axes (the rotation axis, the symmetry axis and the angular momentum axis) and describe the motion of the Earth symmetry poles (*polar motion*) with respect to the instantaneous axis of rotation, which should be really taken as defining the equinox line: the motion thus described includes the so called *polar motion*, *i.e.* the motion of the angular momentum axis (and of the rotation axis) relatively to the symmetry axis. But of course the calculations become intricate and in the end they only provide formulae with free parameters that are determined empirically and used, as said above, for the preparation of the Ephemerides. The nutation motion is simply described by a motion of the Earth symmetry axis endpoint on an ellipse only if the very largest terms from the Moon contributions are considered: all the remaining corrections have the consequence that the motion of the pole around the mean pole is a quasi periodic motion with many periods (ranging from periods of the order of the day up, if one starts including in a very refined theory effects like the tides influence). But the quasi periodic motion can be a good approximation only as long as perturbation theory remains meaningful: the long time behaviour is possibly non quasi periodic and chaotic, even assuming the Earth perfectly rigid: but the chaoticity takes places over a very small scale as the corrections to the main nutational terms correspond to motions of the poles on the Earth surface of the order of $10\,m$ (and the main nutational terms correspond to the order of $100\,m$).

6

Appendices

6.1 A: The Cauchy-Schwartz Inequality

1 Proposition. *Let Ω be a closed bounded Riemann-measurable (or Lebesgue-measurable) set in $\mathcal{R}^d$. Let $f, g \in C^{(0)}(\Omega)$ be two $\mathcal{R}$-valued functions. Then*

$$|\int_\Omega f(\xi)g(\xi)d\xi| \leq \left(\int_\Omega f(\xi)^2 d\xi\right)^{\frac{1}{2}} \left(\int_\Omega g(\xi)^2 d\xi\right)^{\frac{1}{2}}. \qquad (A1)$$

PROOF. In fact $\forall\, \lambda \in \mathcal{R}$:

$$0 \leq \int_\Omega (f(\xi) + \lambda g(\xi))^2 d\xi = \int_\Omega f(\xi)^2 d\xi + 2\lambda \int_\Omega f(\xi)g(\xi)d\xi + \lambda^2 \int_\Omega g(\xi)^2 d\xi.$$

Hence, this polynomial of second degree in λ must have a non-negative discriminant. Its discriminant is simply the difference between the square of the r.h.s. of Eq. (A1) and the square of the l.h.s. mbe

Exercise.

Prove Eq. (A1) by remarking that (if volΔ is the volume of the set Δ)

$$\int_\Omega f(\boldsymbol{\xi})g(\boldsymbol{\xi})d\boldsymbol{\xi} = \lim_{\delta\to 0} \sum_i f(\boldsymbol{\xi}_i)g(\boldsymbol{\xi}_i)\mathrm{vol}\Delta_i$$

where $\Delta_1, \Delta_2, \ldots$ are a pavement of Ω with parallel cubes with side δ and $\boldsymbol{\xi}_i \in \Delta_i \cap \Omega$. Then apply the "ordinary Cauchy inequality" $\sum_i |a_i b_i| \leq (\sum_i |a_i|^2)^{\frac{1}{2}}(\sum_i |b_i|^2)^{\frac{1}{2}}$ to the sequences $a_i = f(\boldsymbol{\xi}_i)\sqrt{\mathrm{vol}\Delta_i}$, $b_i = g(\boldsymbol{\xi}_i)\sqrt{\mathrm{vol}\Delta_i}$.

6.2 B: The Lagrange-Taylor Expansion

1 Proposition. *Let $f \in C^{(k)}(\mathcal{R}^d)$ and suppose that f has a zero of order $(m+1) \leq k$ in $\mathbf{x}_0$. Then*

$$f(\mathbf{x}) = \sum_{\substack{\alpha_1,\ldots,\alpha_d \\ \alpha_i \geq 0, \sum \alpha_i = m+1}} \widetilde{f}_{\mathbf{x}_0,\alpha_1,\ldots,\alpha_d}(\mathbf{x}) \prod_{i=1}^{d} \frac{(x_i - x_{0i})^{\alpha_i}}{a_i!} \tag{B1}$$

and the functions $\widetilde{f}_{\mathbf{x}_0,\alpha_1,\ldots,\alpha_d}(\mathbf{x}) \in C^{(k-m-1}(\mathcal{R}^d)$. If they are regarded as functions of $(\mathbf{x}_0, \mathbf{x}) \in \mathcal{R}^{2d}$ then they are in $C^{(k-m-1)}(\mathcal{R}^{2d})$.

Proof. Consider the function $\lambda \to f(\mathbf{x}_0 + \lambda(\mathbf{x} - \mathbf{x}_0))$ which has in $\lambda = 0$ a zero of order $m+1$, i.e., it has the first m derivatives vanishing. Then

$$\begin{aligned} f(\mathbf{x}) &= \int_0^1 d\lambda_1 \frac{d}{d\lambda_1} f(\mathbf{x}_0 + \lambda_1 (\mathbf{x} - \mathbf{x}_0)) \\ &= \int_0^1 d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \frac{d^2}{d\lambda_2^2} f(\mathbf{x}_0 + \lambda_2 (\mathbf{x} - \mathbf{x}_0)) \\ &= \int_0^1 d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \ldots \int_0^{\lambda_m} d\lambda_{m+1} \frac{d^{m+1}}{d\lambda_{m+1}^{m+1}} f(\mathbf{x}_0 + \lambda_{m+1} (\mathbf{x} - \mathbf{x}_0)) \\ &= \int_0^1 \frac{(1-\lambda)^m}{m!} \frac{d^{m+1}}{d\lambda^{m+1}} f(\mathbf{x}_0 + \lambda (\mathbf{x} - \mathbf{x}_0)) \end{aligned} \tag{B2}$$

Expressing the derivative with respect to λ in terms of the derivatives with respect to the x coordinates it follows, inductively

$$\begin{aligned} &\frac{1}{(m+1)!} \frac{d^{m+1}}{d\lambda^{m+1}} f(\mathbf{x}_0 + \lambda (\mathbf{x} - \mathbf{x}_0)) \\ &= \sum_{\substack{\alpha_1,\ldots,\alpha_d \\ \alpha_i \geq 0, \sum \alpha_i = m+1}} \frac{\partial^{m+1} f(\mathbf{x}_0 + \lambda (\mathbf{x} - \mathbf{x}_0))}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \prod_{i=1}^{d} \frac{(x_i - x_{0i})^{\alpha_i}}{a_i!} \end{aligned} \tag{B3}$$

and this proves the proposition, showing that

$$\widetilde{f}_{\mathbf{x}_0,\alpha_1,\ldots,\alpha_d}(\mathbf{x}) = \int_0^1 \frac{(m+1)!}{m!} \lambda^m \frac{\partial^{m+1} f(\mathbf{x}_0 + \lambda (\mathbf{x} - \mathbf{x}_0))}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \tag{B4}$$

mbe

Observation. The same proof holds if $f \in C^{(k)}(\Omega)$ and Ω is a convex open set.

2 Corollary. *If $f \in C^{(k)}(\mathcal{R}^d \times \mathcal{R}^n)$ has a zero of order $m+1$, $m < k$, in $\mathbf{x}_0$ for each $\mathbf{y} \in \mathcal{R}^n$:*

$$f(\mathbf{x},\mathbf{y}) = \sum_{\substack{\alpha_1,\ldots,\alpha_d \\ \alpha_i\geq 0, \sum \alpha_i = m+1}} \widetilde{f}_{\mathbf{x}_0,\alpha_1,\ldots,\alpha_d}(\mathbf{x},\mathbf{y}) \prod_{i=1}^{d} \frac{(x_i - x_{0i})^{\alpha_i}}{a_i!} \qquad (B5)$$

and the functions $\widetilde{f}$*, thought of as functions of* $(\mathbf{x}_0,\mathbf{x},\mathbf{y}) \in \mathcal{R}^d \times \mathcal{R}^d \times \mathcal{R}^n$*, are in* $C^{(k-(m+1))}(\mathcal{R}^d \times \mathcal{R}^d \times \mathcal{R}^n)$.

Proof. It is a repetition of the above proof.

3 Proposition. *If* $f \in C^{(k)}(\mathcal{R}^d \times \mathcal{R}^n)$ *, the function*

$$f(\mathbf{x},\mathbf{y}) - \sum_{\substack{\alpha_1,\ldots,\alpha_d \\ \alpha_i\geq 0, \sum \alpha_i = \leq m}} \frac{\partial^{\alpha_1+\ldots+\alpha_d} f(\mathbf{x}_0,\mathbf{y})}{\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}} \prod_{i=1}^{d} \frac{(x_i - x_{0i})^{\alpha_i}}{a_i!} \qquad (B6)$$

has, for each $m < k$*, a zero in* $\mathbf{x}$ *of order* m *at* $\mathbf{x}_0$ *for all* $\mathbf{y} \in \mathcal{R}^n$*. Furthermore, the function in Eq. (B6) has a representation like the right hand side of Eq. (B5) with functions* f *having the same properties as those of Corollary 2 above.*

Proof. One first checks that Eq. (B6) has all the $\mathbf{x}$ derivatives vanishing in $(\mathbf{x}_0,\mathbf{y})$ up to order m. Then one applies Corollary 2 or repeats the proof of Proposition 1. This time,

$$\widetilde{f}_{\mathbf{x}_0,\alpha_1,\ldots,\alpha_d}(\mathbf{x},\mathbf{y}) = \int_0^1 \frac{(m+1)!}{m!}\lambda^m \frac{\partial^{m+1} f(\mathbf{x}_0 + \lambda\,(\mathbf{x}-\mathbf{x}_0),\mathbf{y})}{\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}} \qquad (B7)$$

mbe

6.3 C : C^∞-Functions with Bounded Support and Related Functions

1. There is a nonzero function $\psi_a \in C^\infty(\mathcal{R})$, $\psi_a \geq 0$ with support in $[0,a]$ $a > 0$, and one can take

$$\begin{aligned} \psi_a(t) &= 0 && \text{if } t \notin (0,a) \\ \psi_a(t) &= e^{-\frac{1}{t^2(a-t)^2}} && \text{if } t \in (0,a) \end{aligned} \qquad (C1)$$

2. There exists a nondecreasing function $g \in C^\infty(\mathcal{R})$ vanishing for $t \leq 0$ and equal to 1 for $t \geq a > 0$. For instance,

$$\chi(t) = C\int_{-\infty}^{t} \psi_a(\tau)d\tau, \qquad C^{-1} = \int_{-\infty}^{+\infty} \psi_a(\tau)d\tau. \qquad (C2)$$

3. The function

$$g_{\alpha,\beta}(t) = \chi(t-\alpha+a)\chi(\beta-t+a) \tag{C3}$$

has value 1 if $t \in [\alpha,\beta]$, 0 if $t \le \alpha - a$ or $t \ge \beta + a$ and is non-negative.

4. The function in $C^\infty(\mathcal{R}^d)$,

$$g(\xi_1,\ldots,\xi_d) = \prod_{i=1}^{d} g_{\alpha_i,\beta_i}(\xi_i), \tag{C4}$$

has value 1 on the parallelepiped $[\alpha_1,\beta_1] \times \ldots \times [\alpha_d,\beta_d]$; it is 0 outside $[\alpha_1 - a, \beta_1 + a] \times \ldots \times [\alpha_d - a, \beta_d + a]$ and it is non negative.

6.4 D: Principle of the Vanishing Integrals

1 Proposition. *Lt $f \in C^\infty([\alpha,\beta])$ and suppose*

$$\int_\alpha^\beta f(t)z(t)dt = 0 \tag{D1}$$

for all $z \in C_0^\infty([\alpha,\beta])$. Then $f \equiv 0$.

Proof. If $f \not\equiv 0$, there is $t_0 \in (\alpha,\beta)$ where $f(t_0) \neq 0$. Let $[\overline{\alpha},\overline{\beta}] \subset [\alpha,\beta]$ be an interval around t_0 such that $|f(t)| > \frac{1}{2}|f(t_0)|$, $\forall t \in [\overline{\alpha},\overline{\beta}]$. Let $t \to \chi(t)$, $t \in [\alpha,\beta]$ be a C^∞ function positive in t_0 and vanishing outside $[\overline{\alpha},\overline{\beta}]$. Then $t \to f(t_0)\chi(t)$ is in $C_0^\infty([\alpha,\beta])$ and

$$0 = \int_\alpha^\beta f(t)\chi(t)f(t_0)dt \ge \frac{1}{2}\int_{\overline{\alpha}}^{\overline{\beta}} |f(t_0)|^2\chi(t)f(t_0)dt > 0. \tag{D2}$$

mbe

6.5 E: Matrix Notations. Eigenvalues and Eigenvectors. A List of some Basic Results in Algebra

The reader who wishes more details (or proofs) on the subjects discussed below may consult [14] Chaps. 1 and 2.

1. Given a $\ell\times m$ matrix J, and a $m\times p$ matrix L, JL denotes the $\ell\times p$ matrix obtained by multiplying "rows by columns" the matrices J and L.

2. If J is an $\ell \times m$ matrix and $\mathbf{x} \in \mathcal{C}^m$, we denote $\mathbf{y} = L\mathbf{x}$ the vector of $\mathcal{C}^\ell$ with components

$$y_i = \sum_{k=1}^{m} J_{ik} x_k, \qquad i = 1, \dots, \ell. \tag{$E1$}$$

3. The determinant, $\det J$, of a matrix J is defined for all the square matrices. If J and L are two $d \times d$ square matrices, $\det JL = \det J \det L$. The determinant is a linear combination of products of matrix elements.

4. The sum of two $\ell \times m$ matrices is an $\ell \times m$ matrix with matrix elements given by the sums of the homonymous matrix elements of the two matrices. The matrix $\lambda J, \lambda \in C$, is the matrix whose elements are those of J multiplied by λ. The modulus of an $\ell \times m$ matrix J is

$$|J| \stackrel{def}{=} \sum_{i=1}^{\ell} \sum_{j=1}^{m} |J_{ij}|. \tag{$E2$}$$

If J is an $\ell \times m$ matrix and L is a $m \times p$ matrix,

$$|JL| \le |J|\,|L|. \tag{$E3$}$$

In §5.12 (only) we use the symbol $||J||$ for the right-hand side of (E2) and $|J|$ for $\max|J_{ij}|$; then Eq. (E3) is changed by an extra factor ℓm in the right hand side.

5. The $d\times d$ identity matrix, will usually be simply denoted by 1 and similarly, the product of $\lambda \in \mathcal{C}$ with the identity matrix will be denoted λ.

6. The eigenvalues of a square matrix are the solutions of the algebraic equation in λ ("secular or characteristic equation"):

$$\det(J - \lambda) = 0. \tag{$E4$}$$

7. The inverse matrix to a square matrix J exists if and only if $\det J \neq 0$ and it will be denoted J^{-1}: it is characterized by the property $JJ^{-1} = J^{-1}J = 1$. Its matrix elements are expressible as ratios of determinants of submatrices of J by the determinant of J.

8. If $f(z) = \sum_{n=0}^{\infty} c_n z^n$ is a power series with radius of convergence $\varrho > 0$ and if J is a square matrix such that $|J| < \varrho$ and if $J^0 \stackrel{def}{=} 1$, the series

$$f_{ij} = \sum_{n=0}^{\infty} c_n (J^n)_{ij} \tag{E5}$$

are absolutely convergent since

$$|f_{ij}| = \sum_{n=0}^{\infty} |c_n||J^n| \le \sum_{n=0}^{\infty} |c_n||J|^n \le \sum_{n=0}^{\infty} |c_n|\varrho^n < \infty \tag{E6}$$

They define a matrix that will be denoted $f(J)$.
If $(P(z), Q(z))$ are two polynomials and $PQ(z)$ is their product polynomial, it is

$$P(J)Q(J) = PQ(J) \tag{E7}$$

(if one thinks of the definition of the product of polynomials and of the fact that the product of matrices is distributive).
Similarly, if $f(z), g(z)$ are two powers series with radius of convergence ϱ, their product power series $fg(z)$ has the same radius of convergence and the above relation is generalized by

$$f(J)g(J) = fg(J). \tag{E8}$$

In particular, if $|J| < 1$, $f(z) = 1 - z$, $g(z) = (1-z)^{-1} = \sum_{n=0}^{\infty} z^n$, $fg(z) = 1$, so that $g(J)$ is the inverse to $(1 - J)$; i.e.,

$$(1 - J)^{-1} = \sum_{n=0}^{\infty} J^n \tag{E9}$$

and

$$|(1 - J)^{-1} - 1| = |\sum_{n=1}^{\infty} J^n| \le \frac{|J|}{1 - |J|} \tag{E10}$$

9. A real square matrix J is said to be "orthogonal" if $J^{-1} = J^T$, where $(J^T)_{ij} \stackrel{def}{=} J_{ji}$, $\forall\, i, j$. The orthogonal $d \times d$ matrices can also be thought of as "rotations of $\mathcal{R}^d$". The rotation of $\mathcal{R}^d$ corresponding to the orthogonal matrix J will be the map of $\mathcal{R}^d$ into itself:

$$\mathbf{x} \to J\mathbf{x} \tag{E11}$$

10. If J is a $d \times d$ matrix and if $\mathbf{y}, \mathbf{x} \in \mathcal{C}^d$,

$$\mathbf{x} \cdot J\mathbf{y} = J^T\mathbf{x} \cdot \mathbf{y}. \tag{E12}$$

11. The eigenvalues of a matrix enjoy remarkable properties. For instance:

1 Proposition. *If J is a $d \times d$ matrix with pairwise-distinct eigenvalues $\lambda_1, \ldots, \lambda_d$, there are d vectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)} \in \mathcal{C}^d$, generally complex even if J is a real matrix, such that*

$$J\mathbf{v}(i) = \lambda_i \mathbf{v}^{(i)}, \qquad i = 1, \ldots, d \qquad (E13)$$

and they are linearly independent.
If J is a real matrix, the eigenvalues and the eigenvectors can be arranged so that they appear in complex-conjugate pairs.
If the matrix J varies in the neighborhood (in the sense that $|J - J_0|$ is small) of a matrix J_0 with pairwise-distinct eigenvalues, then the eigenvalues and the corresponding eigenvectors can be chosen and labeled so that they vary smoothly with J, i.e., so that the eigenvalue λ_j of J and the corresponding eigenvector components $(\mathbf{v}^{(j)})_k$, $j, k = 1, \ldots, d$, are C^∞ functions of the matrix elements of J.

6.6 F: Positive-Definite Matrices. Eigenvalues and Eigenvectors. A List of Basic Properties

The reader who wishes more details (or proofs) on the subjects discussed below may consult [14], Chaps. 1 and 2.

Definition. *A real matrix $V = (V_{ij})$, $i, j = 1, \ldots, d$ is "positive definite" if*

$$(i) \qquad V_{ij} = V_{ji},\ i, j = 1, \ldots, d; \qquad \text{(Symmetry)} \qquad (F1)$$

(ii) For all $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathcal{R}^d$, $\boldsymbol{\alpha} \neq \mathbf{0}$,

$$(\boldsymbol{\alpha} \cdot \boldsymbol{\alpha}) \stackrel{def}{=} \sum_{i,j=1}^{d} V_{ij} \alpha_i \alpha_j > 0 \qquad \text{(Positivity)} \qquad (F2)$$

We now collect the main properties of the positive-definite matrices in two propositions.

First, note that $\det V \neq 0$; otherwise, there would be $\boldsymbol{\alpha}_0 \neq \mathbf{0}$ such that $V\boldsymbol{\alpha}_0 = \mathbf{0}$, contradicting (ii) above.

The following proposition states the "existence of an orthonormal basis on which V is diagonal".

1 Proposition. *If V is a $d \times d$ positive-definite matrix, there exist d positive numbers $\lambda_1, \ldots, \lambda_d$ and an orthonormal basis $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}$ in $\mathcal{R}^d$ such that*

$$V\mathbf{v}^{(j)} = \lambda_j \mathbf{v}^{(j)}, \qquad j = 1, \ldots, d \qquad (F3)$$

and the orthogonal matrix

$$J_{ij} = (\mathbf{v}^{(j)})_i, \qquad i, j = 1, \ldots, d \tag{F4}$$

is such that

$$JVJ^T = \Lambda, \qquad V = J^T \Lambda J, \tag{F5}$$

where Λ is the diagonal $d\times d$ matrix with diagonal elements given by $\lambda_1, \ldots, \lambda_d$.

Observation. Eq. (F5) implies that $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of V counted according to multiplicity. In fact,

$$\begin{aligned}\det(V-\lambda) =& (\det J^T\Lambda J - \lambda) = \det(J^T(\Lambda-\lambda)J)\\ =& \det(\Lambda-\lambda) = \prod_{i=1}^{d}(\lambda_i - \lambda)\end{aligned} \tag{F6}$$

(since $J^TJ \equiv 1$, $\det J\ \det J^T = \det JJ^T = 1$)

2 Corollary. *If V is a positive-definite $d\times d$ matrix, there is a positive definite matrix $\sqrt{V}$ such that $(\sqrt{V})^2 = V$. More generally, if $a \in \mathcal{R}$, there is a positive-definite matrix V^a such that, $\forall\, a, b \in \mathcal{R}$, $V^aV^b = V^{a+b}$ and $V^1 = V$, $V^0 = 1$.*

Proof. If Λ is a diagonal $d \times d$ matrix such that Eq. (F5) holds, we set $\Lambda^a = \{$diagonal matrix and diagonal elements $\lambda_1, \ldots, \lambda_d\}$. Then $\Lambda^a\Lambda^b = \Lambda^{a+b}$, $\forall\, a, b \in \mathcal{R}, \Lambda^1 = \Lambda, \Lambda^0 = 1$; so we set

$$V^a = J^T \Lambda^a J \tag{F7}$$

and V^a verifies the desired properties. $V^{\frac{1}{2}} = \sqrt{V}$ by definition. mbe

3 Corollary. *If V is a $d\times d$ positive-definite matrix, there exists a continuous function $\mu(V) > 0$ depending on the matrix elements of V such that*

$$V\boldsymbol{\alpha}\cdot\boldsymbol{\alpha} \geq \mu(V)|\boldsymbol{\alpha}|^2. \tag{F8}$$

In fact, $\mu(V) = \min_i \lambda_i$, are the eigenvalues of V.

Proof.

$$\begin{aligned}V\boldsymbol{\alpha}\cdot\boldsymbol{\alpha} =& J^T\Lambda J\boldsymbol{\alpha}\cdot\boldsymbol{\alpha} = \Lambda J\boldsymbol{\alpha}\cdot J\boldsymbol{\alpha} = \sum_{i=1}^{d}\lambda_i(JBa)_i^2\\ \geq& \mu(V) = \sum_{i=1}^{d}(JBa)_i^2 = \mu(V)\, J\boldsymbol{\alpha}\cdot J\boldsymbol{\alpha}\\ =& \mu(V)\, J^TJ\boldsymbol{\alpha}\cdot\boldsymbol{\alpha} = \mu(V)\,\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}.\end{aligned} \tag{F9}$$

mbe

We conclude with a generalization of the above results.

4 Proposition. *Let G, V be two positive-definite $d \times d$ matrices. There are d independent vectors $\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)} \in \mathcal{R}^d$ and d positive numbers $\lambda_1, \ldots, \lambda_d$ such that*

$$V\mathbf{v}^{(j)} = \lambda_j G\mathbf{v}^{(j)}, \qquad j = 1, \ldots, d, \tag{F10}$$

$$G\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)} = \delta_{ij}, \qquad i, j = 1, \ldots, d; \tag{F11}$$

The numbers $\lambda_1, \ldots, \lambda_d$ are the solutions repeated with multiplicity of

$$\det(V - \lambda G) = 0. \tag{F12}$$

There is a function $\mu(V, G) > 0$ continuously dependent on the matrix elements of V, G such that

$$V\boldsymbol{\alpha} \cdot \boldsymbol{\alpha} \geq \mu(V, G)\,(G\boldsymbol{\alpha} \cdot \boldsymbol{\alpha}). \tag{F13}$$

Observation. This is reduced to the preceding propositions. If $\mathbf{w}^{(1)}, \ldots, \mathbf{w}^{(d)}$ are the eigenvectors of the positive-definite matrix $W = G^{-\frac{1}{2}} V G^{-\frac{1}{2}}$, the $\mathbf{v}^{(j)}$ are

$$\mathbf{v}^{(j)} = G^{-\frac{1}{2}}\mathbf{w}^{(j)}, \qquad j = 1, \ldots, d. \tag{F14}$$

6.7 G: Implicit Functions Theorems

Let $\mathbf{f} \in C^\infty(\mathcal{R}^m \times \mathcal{R}^d)$ be a function with values in $\mathcal{R}^d$ associating to $(\mathbf{x}, \mathbf{y}) \in \mathcal{R}^m \times \mathcal{R}^d$ the value $\mathbf{f}(\mathbf{x}, \mathbf{y})$.

Consider the equation for $\mathbf{y} \in \mathcal{R}^d$ parameterized by $\mathbf{x}$:

$$\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0} \tag{G1}$$

which is a system with d equations in d unknowns $y_1, \ldots, y_d$.

Suppose that $(\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{R}^m \times \mathcal{R}^d$ verifies Eq. (G1). By the Taylor theorem, see Appendix B,

$$\mathbf{f}(\mathbf{x}, \mathbf{y}) = J\,(\mathbf{y} - \mathbf{y}_0) + L\,(\mathbf{x} - \mathbf{x}_0) + \mathbf{N}(\mathbf{x}, \mathbf{y}) \tag{G2}$$

where J, L are $d \times d$ matrices built with the derivatives of f:

$$J_{ij} = \frac{\partial f^{(i)}}{\partial y_j}(\mathbf{x}_0, \mathbf{y}_0), \qquad i, j = 1, \ldots, d, \tag{G3}$$

$$L_{ij} = \frac{\partial f^{(i)}}{\partial x_j}(\mathbf{x}_0, \mathbf{y}_0), \qquad i = 1, \ldots, d,\ j = 1, \ldots, m \tag{G4}$$

and $\mathbf{N}$ is an $\mathcal{R}^d$-valued C^∞-function with a second-order zero in $(\mathbf{x}_0, \mathbf{y}_0)$, see Appendix B, Proposition 3 with $m = 1, k = +\infty$

The implicit function theorem compares the solution of Eq. (G1), written as

$$J\,(\mathbf{y} - \mathbf{y}_0) + L\,(\mathbf{x} - \mathbf{x}_0) + \mathbf{N}(\mathbf{x}, \mathbf{y}) = \mathbf{0}, \tag{G5}$$

with that of the linear equations ($d \times d$ linear system)

$$J\,(\mathbf{y} - \mathbf{y}_0) + L\,(\mathbf{x} - \mathbf{x}_0) = \mathbf{0}. \tag{G6}$$

If $\det J \neq 0$, the matrix J^{-1} exists and Eq. (G6) has the unique solution

$$\mathbf{y} - \mathbf{y}_0 = -J^{-1}\,L\,(\mathbf{x} - \mathbf{x}_0). \tag{G7}$$

Therefore, it becomes natural to think that Eq. (G5) admits a solution differing from Eq. (G7) "by higher-order infinitesimals in $\mathbf{x} - \mathbf{x}_0$", since such is the difference between Eq. (G5) and Eq. (G6). More precisely, one can hope that there exists in a vicinity U of $\mathbf{x}_0$ a function $\boldsymbol{\varphi}(\mathbf{x})$ such that

$$\mathbf{f}(\mathbf{x}, \boldsymbol{\varphi}(\mathbf{x})) \equiv \mathbf{0}, \qquad \mathbf{x} \in U, \tag{G8}$$

$$\boldsymbol{\varphi}(\mathbf{x}) = -J^{-1}\,L\,(\mathbf{x} - \mathbf{x}_0) + \boldsymbol{\Phi}(\mathbf{x}), \qquad \mathbf{x} \in U \tag{G9}$$

where $\boldsymbol{\Phi} \in C^\infty(U)$ and has a second-order zero at $\mathbf{x}_0$.

This is, in fact, the content of the implicit function theorems. Since we shall also need explicit estimates of the size of the set U, on which $\boldsymbol{\Phi}$ can be defined, and on the size of $\boldsymbol{\Phi}(U)$, its $\boldsymbol{\Phi}$ image, it is more appropriate to describe the proof in notations which are convenient for us rather than to refer to a standard book.

We first treat the $d = 1$ case, denoting $\Gamma_n(\mathbf{x}, \varrho) \subset \mathcal{R}^n$ the closed cube with center $\mathbf{x} \in \mathcal{R}^n$ and side 2ϱ.

1 Proposition. *Given $\delta > 0, \alpha > 0, \alpha \geq \delta$, define*

$$\varrho_{\delta,\alpha} \stackrel{def}{=} \frac{\delta}{2} \frac{\min |\frac{\partial f}{\partial y}|}{\max(\sum_{j=1}^m |\frac{\partial f}{\partial x_j}| + |\frac{\partial f}{\partial y}|)}, \tag{G10}$$

where the minimum and the maximum are considered as $\mathbf{x}$ varies in $\Gamma_m(\mathbf{x}_0, \alpha)$ and as $y - y_0$ varies in $[-\delta, \delta]$ and we suppose that $f(\mathbf{x}_0, y_0) = 0$.
If $\varrho_{\delta,\alpha} > 0$ it is possible to define a function φ in $C^\infty(\Gamma_m(\mathbf{x}_0, \varrho_{\delta.\alpha}))$ verifying Eq. (G8) for every $\mathbf{x} \in \Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$.
Furthermore, all the solutions of Eq. (G1) in $\Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha}) \times [y_0 - \delta, y_0 + \delta]$ have the form $(\mathbf{x}, \varphi(\mathbf{x}))$ and

$$\frac{\partial \varphi}{\partial x_i}(\mathbf{x}) = -\frac{\frac{\partial f}{\partial x_i}(\mathbf{x}, \varphi(\mathbf{x}))}{\frac{\partial f}{\partial f}(\mathbf{x}, \varphi(\mathbf{x}))}. \tag{G11}$$

PROOF. Let $(\mathbf{x}, y_0+\delta)$ be a point on the upper face of the parallelepiped $\Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha}) \times [y_0-\delta, y_0+\delta]$. We show that on this face f has a well-defined sign, *opposite* to the one it has on the lower face.

Since $\frac{\partial f}{\partial y} \neq 0$ cannot vanish in the parallelepiped, by the choice of $\varrho_{\delta,\alpha}$ and because $\varrho_{\delta,\alpha} > 0$ this will imply that for each $\mathbf{x} \in \Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$ there is and only one point $\varphi(\mathbf{x}) \in [y_0-\delta, y_0+\delta]$ such that $f(\mathbf{x}, \varphi(\mathbf{x})) = 0$ (note as that $\frac{\partial f}{\partial y} \neq 0$ implies strict monotonicity).

To show that f takes opposite signs on the opposite faces suppose, to be definite, $\frac{\partial f}{\partial y} > 0$ in $\Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha}) \times [y_0-\delta, y_0+\delta]$. Then

$$\begin{aligned}\mathbf{f}(\mathbf{x}, y+\delta) \equiv& f(\mathbf{x}, y_0+\delta) - f(\mathbf{x}_0, y_0)\\ =& f(\mathbf{x}, y_0+\delta) - f(\mathbf{x}, y_0) + f(\mathbf{x}, y_0+\delta) - f(\mathbf{x}_0, y_0)\end{aligned} \tag{G12}$$

and we apply the Lagrange theorem to find $\widetilde{\mathbf{x}}$ and $\widetilde{y}$, intermediate between $\mathbf{x}$ and $\mathbf{x}_0$ and between y_0 and $y_0+\delta$, such that the right-hand side of Eq. (G12) can be written

$$\begin{aligned}f(\mathbf{x}, y_0+\delta) =& \frac{\partial f}{\partial y}(\mathbf{x}, \widetilde{y})\,\delta + \sum_{j=1}^m \frac{\partial f}{\partial x_j}(\widetilde{\mathbf{x}}, y_0)(\mathbf{x}-\mathbf{x}_0)\\ \geq& (\min|\frac{\partial f}{\partial y}|)\,\delta - (\max\sum_{j=1}^m |\frac{\partial f}{\partial x_j}|)\,\varrho_{\delta,\alpha}\end{aligned} \tag{G13}$$

Similarly, one proves that $f(\mathbf{x}, y_0-\delta) < 0, \forall\, \mathbf{x} \in \Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$. This proves the existence of $\beta(\mathbf{x})$ and its uniqueness.

To show the differentiability in the direction of the axis $\mathbf{e} = (e_1, \ldots, e_d)$ observe that, given $\mathbf{x} \in \Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$ and given $\mathbf{e}$ such that $\mathbf{x}_0+\varepsilon\,\mathbf{e} \in \Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$ and if $\widetilde{\mathbf{x}}, \widetilde{y}$ are suitable intermediate points between $\mathbf{x}$ and $\mathbf{x}+\varepsilon\mathbf{e}$ or $\varphi(\mathbf{x})$ and $\varphi(\mathbf{x}+\varepsilon\,\mathbf{e})$, one finds

$$0 \equiv f(\mathbf{x}+\varepsilon\mathbf{e}, \varphi(\mathbf{x}+\varepsilon\mathbf{e})) - \varphi(\mathbf{x}, \varphi(\mathbf{x})), \tag{G14}$$

$$0 = \sum_{i=1}^m e_i \frac{\partial f}{\partial x_i}(\widetilde{\mathbf{x}}, \widetilde{y})\,\varepsilon + \frac{\partial f}{\partial y}(\widetilde{\mathbf{x}}, \widetilde{y})(\varphi(\mathbf{x}+\varepsilon\mathbf{e}) - \varphi(\mathbf{x}) \tag{G15}$$

by the Lagrange theorem, and this shows that

$$|\varphi(\mathbf{x}+\varepsilon\mathbf{e}) - \varphi(\mathbf{x}) \leq \frac{\max\sum_{i=1}^m |\frac{\partial f}{\partial x_i}|}{\min|\frac{\partial f}{\partial y}|} \tag{G16}$$

i.e., φ is continuous. Eq. (G15) also yields, dividing it by ε and letting $\varepsilon \to 0$,

$$\lim_{\varepsilon\to 0} \frac{\varphi(\mathbf{x}+\varepsilon\mathbf{e}) - \varphi(\mathbf{x}, \varphi(\mathbf{x}))}{\varepsilon} = \frac{-\sum_{i=1}^m e_i\,\frac{\partial f}{\partial x_i}(\mathbf{x}, \varphi(\mathbf{x}))}{\frac{\partial f}{\partial y}(\mathbf{x}, \varphi(\mathbf{x}))}, \tag{G17}$$

proving the differentiability of φ and Eq. (G11).

By the chain-differentiation rule of composed-function, Eq. (G11) implies that $\frac{\partial\varphi}{\partial x_j}$ are differentiable in $\mathbf{x}$ and their derivatives can be expressed in terms of φ, of its first derivative and of f and its first two partial derivatives. Therefore, $\frac{\partial^2\varphi}{\partial x_i \partial x_j}$ are differentiable, etc., i.e., $\varphi \in C^\infty(\Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$. mbe

Observation. It appears from the proof that the same results hold if no relation is assumed a priori between α and δ provided $\varrho_{\delta,\alpha}$ replaced by

$$\overline{\varrho}_{\delta,\alpha} = \Big\{ \text{minimum between } \alpha \text{ and } \tfrac{1}{2}\delta \frac{\min|\frac{\partial f}{\partial y}|}{\max\sum_{i=1}^m |\frac{\partial f}{\partial x_i}|} \Big\} \tag{G18}$$

which is a better result; see Eq. (G13).

To deal with the general case, introduce, given a matrix M,

$$|M| = \sum_{i,j} |M_{ij}| \tag{G19}$$

and note that $|M \cdot N| \le |M|\,|N|$ if $M \cdot N$ makes sense, i.e., if the number of columns of M equals that of the rows of N. Also define the matrices

$$J(\mathbf{x},\mathbf{y}) \stackrel{def}{=} \frac{\partial \mathbf{f}}{\partial \mathbf{y}}, \qquad L(\mathbf{x},\mathbf{y}) \stackrel{def}{=} \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \tag{G20}$$

2 Proposition. *Given* $\delta, \alpha > 0$, *define*

$$\varrho_{\delta,\alpha} \stackrel{def}{=} \frac{1}{2} \frac{\delta - 2(\max|J^{-1}|)(\alpha \max|\frac{\partial \mathbf{N}}{\partial \mathbf{x}}| + \delta \max|\frac{\partial \mathbf{N}}{\partial \mathbf{y}}|)}{\max|J^{-1}L|} \tag{G21}$$

with the maxima taken on $\Gamma_m(\mathbf{x}_0, \alpha) \times \Gamma_d(\mathbf{y}_0, \delta)$ *and set* $\varrho_{\delta,\alpha} = 0$ *if* J^{-1} *does not exist at some point of this set. Suppose* $\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$, *and* $\alpha > \varrho_{\alpha,\delta} > 0$. *It is then possible to find* $\boldsymbol{\varphi} \in C^\infty(\Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha}))$ *with values in* $\Gamma_d(\mathbf{y}_0, \delta)$ *verifying Eq. (G8).*
Furthermore, all the solutions of Eq. (G1) in $\Gamma_m(\mathbf{x}_0, \alpha) \times \Gamma_d(\mathbf{y}_0, \delta)$ *have the form* $(\mathbf{x}, \boldsymbol{\varphi}(\mathbf{x}))$ *and*

$$\frac{\partial \varphi_k}{\partial x_j} = -\sum_{k=1}^{d} \Big(\frac{\partial \mathbf{f}(\mathbf{x}, \boldsymbol{\varphi}(x))}{\partial \mathbf{y}}\Big)^{-1}_{ki} \Big(\frac{\partial f_i(\mathbf{x}, \boldsymbol{\varphi}(x))}{\partial x_j}\Big), \qquad \mathbf{x} \in \Gamma(\mathbf{x}_0, \varrho_{\delta,\alpha}) \tag{G22}$$

Observations.
(1) Note that the above proposition is nonempty. Using the fact$\mathbf{N}$ has a second-order zero at $(\mathbf{x}_0, \mathbf{y}_0)$, given $B > |J(\mathbf{x}_0, \mathbf{y}_0)^{-1} JL(\mathbf{x}_0, \mathbf{y}_0)|^{-1}$, we see that for δ small enough (depending on B) and $\alpha = B\delta$ it is:

$$0 < \frac{1}{2} \frac{\delta}{\max|J^{-1}L|} < \varrho_{\delta,\alpha} < \frac{\delta}{\max|J^{-1}L|} < B\delta \tag{G23}$$

(2) There are two methods to prove a theorem like the above. The most natural would be to deduce it as a corollary of Proposition 1. One would just proceed by substitution as in the solution of the linear systems.

The assumption $\det J \neq 0$ implies that there is at least one derivative $\frac{\partial f^{(i_1)}}{\partial y_1}$. Then we apply Proposition 1 to the function $f = f^{(i_1)}$ with $y = y_1$ and $\mathbf{x}$ replaced by $(\mathbf{x}, y_2, \ldots, y_d)$ and call $\varphi_1(\mathbf{x}, y_2, \ldots, y_d)$ its solution defined close enough to $\mathbf{x}_0, y_{0\,2}, y_{0\,3}, \ldots, y_{0\,d}$. Then, supposing $i_1 = 1$, consider

$$\begin{aligned} &f^{(2)}(\mathbf{x}, \varphi_1(\mathbf{x}, y_2, \ldots, y_d), y_2, \ldots, y_d) = 0 \\ &\ldots\ldots \\ &f^{(d)}(\mathbf{x}, \varphi_1(\mathbf{x}, y_2, \ldots, y_d), y_2, \ldots, y_d) = 0 \end{aligned} \tag{G24}$$

The determinant of the Jacobian matrix J, of the left-hand side of Eq. (G24) with respect to $y_2, \ldots, y_d$ cannot vanish in $\mathbf{x}_0, y_{0\,2}, y_{0\,3}, \ldots, y_{0\,d}$ because it can be shown to coincide with the determinant of the linear system of equations obtained from the system $J(\mathbf{x}_0, \mathbf{y}_0)\boldsymbol{\xi} = \boldsymbol{\eta}$ by solving its first equation with respect to ξ_1 and substituting into the others. Therefore, we can again apply Proposition 1, expressing, say, y_2 as a function of $\mathbf{x}, y_2, \ldots, y_d$ close enough to $\mathbf{x}_0, y_{0\,3}, \ldots, y_{0\,d}$ etc. The only difficulty is that the left-hand side of Eq. (G24) is only defined, and C^∞, in a small vicinity of $\mathbf{x}_0, (\mathbf{y}_0)_2, \ldots, (\mathbf{y}_0)_d$, and not on all of $\mathcal{R}^m \times \mathcal{R}^d$, as would be required by Proposition 1. This is, however, an obviously trivial difficulty. What is more difficult in this method is to keep track of the size of the neighborhoods involved, in order to obtain an explicit formula like Eq. (G21). Therefore, here we shall adopt another classical method of proof. The triumph of the naive substitution method will appear in Appendix N where, however, additional assumptions on f are made.

Proof. Write Eq. (G1) as Eq. (G5) and let

$$\mathbf{y}' - \mathbf{y}_0 = -J^{-1}L(\mathbf{x} - \mathbf{x}_0) - J^{-1}\mathbf{N}(\mathbf{x}, \mathbf{y}) \tag{G25}$$

for $(\mathbf{x}, \mathbf{y}) \in \varGamma_m(\mathbf{x}_0, \alpha) \times \varGamma_d(\mathbf{y}_0, \delta)$. Note that $\mathbf{y}' - \mathbf{y}o| < \delta$ if $(\mathbf{x}, \mathbf{y}) \in \varGamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha}) \times \varGamma_d(\mathbf{y}_0, \delta)$ and if (as supposed) $a > \varrho_{\delta,\alpha} > 0$. In fact, by the Lagrange theorem and $\mathbf{N}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$, it follows that

$$\begin{aligned} |\mathbf{y}' - \mathbf{y}_0| \leq & |J^{-1}L|\, \varrho_{\delta,\alpha} + |J^{-1}(\mathbf{N}(\mathbf{x}, \mathbf{y}) - N(\mathbf{x}_0, \mathbf{y}_0))| \\ \leq & |J^{-1}L|\, \varrho_{\delta,\alpha} + max|J^{-1}|\big(|\frac{\partial \mathbf{N}}{\partial \mathbf{x}}|\alpha + |\frac{\partial \mathbf{N}}{\partial \mathbf{y}}|\delta\big) < \frac{1}{2}\delta \end{aligned} \tag{G26}$$

Therefore, at $\mathbf{x}$ fixed in $\varGamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$, Eq. (G25) yields a map of $\varGamma_d(\mathbf{y}_0, \varrho_{\delta,\alpha})$ into itself. We can, therefore, recursively define, for each fixed $\mathbf{x} \in \varGamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha})$,

$$\mathbf{y}_n - \mathbf{y}_0 = -J^{-1}L(\mathbf{x} - \mathbf{x}_0) - J^{-1}\mathbf{N}(\mathbf{x}, \mathbf{y}_{n-1} \tag{G27}$$

$n = 1, 2, \ldots$. Then,

$$|\mathbf{y}_n - \mathbf{y}_{n-1}| = |J^{-1}(\mathbf{N}(\mathbf{x}, \mathbf{y}_{n-1} - \mathbf{N}(\mathbf{x}, \mathbf{y}_{n-2})| \\ \le (\max |J^{-1}|)\big(\max |\frac{\partial \mathbf{N}}{\partial \mathbf{y}}|\big)\, |\mathbf{y}_{n-1} - \mathbf{y}_{n-2}| < \frac{1}{2}|\mathbf{y}_{n-1} - \mathbf{y}_{n-2}| \tag{G28}$$

having used in the last step the hypothesis $\varrho_{\delta,\alpha} > 0$ which implies that $(\max |J^{-1}|)(\max |\frac{\partial \mathbf{N}}{\partial \mathbf{y}}|) < \frac{1}{2}$. Therefore, $|\mathbf{y}_n - \mathbf{y}_{n-1}| \le 2^{-(n-1)}|\mathbf{y}_1 - \mathbf{y}_0|$ and there exists the limit

$$\boldsymbol{\varphi}(\mathbf{x}) = \lim_{n\to\infty} \mathbf{y}_n = \mathbf{y}_0 + \sum_{k=1}^{\infty}(\mathbf{y}_k - \mathbf{y}_{k-1}) \tag{G29}$$

If $(\mathbf{x}, \widetilde{\mathbf{y}})$ is another solution to Eq. (G1) in $\Gamma_m(\mathbf{x}_0, \varrho_{\delta,\alpha}) \times \Gamma_d(\mathbf{y}_0, \delta)$, we can write Eq. (G1) in the form of Eq. (G5) for $\mathbf{y}$ and $\boldsymbol{\varphi}(\mathbf{x})$ and subtract

$$|\widetilde{\mathbf{y}} - \boldsymbol{\varphi}(\mathbf{x})| = |J^{-1}(\mathbf{N}(\mathbf{x}, \widetilde{\mathbf{y}}) - \mathbf{N}n(\mathbf{x}, \boldsymbol{\varphi}(\mathbf{x}))| \le \frac{1}{2}|\widetilde{\mathbf{y}} - \boldsymbol{\varphi}(\mathbf{x})| \tag{G30}$$

i.e. $\widetilde{\mathbf{y}} = \boldsymbol{\varphi}(\mathbf{x})$, proving uniqueness. The differentiability statement is proved as in Proposition 1. mbe

3 Corollary. *Under the assumptions of Proposition 2, let $m = d$ and, see Eq. (G23), give $B, C > 1$ such that*

$$B > (\min |J^{-1}L|)^{-1}, \qquad C > (\min |L^{-1}J|)^{-1}, \tag{G31}$$

where the minima are taken over $\Gamma_d(\mathbf{x}_0, \overline{\alpha}) \times \Gamma_d(\mathbf{y}_0, \overline{\delta})$ with given $\overline{\alpha}, \overline{\delta} > 0$. Suppose that $\delta > 0$ is so small that $\delta, B\delta < \overline{\alpha}, \overline{\delta}$.
Define $\varrho_{\alpha,\delta}$ as in Eq. (G21) and $\widetilde{\varrho}_{\alpha,\delta}$

$$\widetilde{\varrho}_{\alpha,\delta} = \frac{1}{2}\frac{\alpha - 2(\max |L^{-1}|)(\alpha \max |\frac{\partial \mathbf{N}}{\partial \mathbf{x}}| + \delta |\frac{\partial \mathbf{N}}{\partial \mathbf{y}}|)}{\max |L^{-1}J|}, \tag{G32}$$

where the maxima are now considered on $\Gamma_d(\mathbf{x}_0, \overline{\alpha}) \times \Gamma_d(\mathbf{y}_0, \overline{\delta})$ both for $\varrho_{\delta,\alpha}$ and $\widetilde{\varrho}_{\delta,\alpha}$. Then if δ is so small that

$$0 < \varrho \overset{def}{=} \varrho_{\delta, B\delta} < B\delta, \qquad \text{and} \quad 0 < \widetilde{\varrho} \overset{def}{=} \widetilde{\varrho}_{\frac{1}{BC}\varrho, \frac{1}{B}\varrho} < \delta \tag{G33}$$

(which is possible by Observation (1), p.530), the $\boldsymbol{\varphi}$-image of $\Gamma_d(\mathbf{x}_0, \varrho)$ covers $\Gamma_d(\mathbf{y}_0, \widetilde{\varrho})$.

Observations.
(1) This means that if the Jacobians of $\mathbf{f}$ with respect to $\mathbf{x}$ and with respect to $\mathbf{y}$ have non vanishing determinant at $(\mathbf{x}_0, \mathbf{y}_0)$, the $\mathbf{f}$ sets up a correspondence between $\mathbf{x}, \mathbf{y}$ near $\mathbf{x}_0, \mathbf{y}_0$ of class C^∞, with inverse of class C^∞, and sending open sets onto open sets (it is a local "C^∞ diffeomorphism").
(2) Since Corollary 3 is quantitative, it says much more: it gives, in fact, estimates of the size of the regions where $\mathbf{f}$ can be inverted.

Proof. Just apply Proposition 2 twice, to express $\mathbf{y}$ in terms of $\mathbf{x}$ and viceversa (make a two-dimensional drawing to better understand the situation). mbe

Another important application of Proposition 2 is the following corollary used in §5.10.

4 Corollary. *Let* $\mathbf{f} \in C^\infty(\mathcal{T}^\ell)$ *with values in* $\mathcal{R}$. *Consider the equation for* $\boldsymbol{\varphi} \in \mathcal{T}^\ell$:

$$\boldsymbol{\varphi}' = \boldsymbol{\varphi} + \varepsilon\, f(\boldsymbol{\varphi}) \tag{G34}$$

with $\varepsilon \in \mathcal{R}_+$ *and suppose* $\max|\mathbf{f}(\boldsymbol{\varphi})| \le 1$, $\max|\frac{\partial \mathbf{f}}{\partial \boldsymbol{\varphi}}(\boldsymbol{\varphi})| \le 1$.
There is $\varepsilon_\ell > 0$, *depending only on* ℓ *and not on* $\mathbf{f}$, *such that,* $\forall\, \varepsilon < \varepsilon_\ell$, *the above equation can be solved uniquely in the form*

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}' + \varepsilon\, \mathbf{g}(\boldsymbol{\varphi}', \varepsilon) \tag{G35}$$

with $\mathbf{g} \in C^\infty(\mathcal{T}^\ell)$ *at fixed* ε. *Furthermore* $\max_{\boldsymbol{\varphi}} |\mathbf{g}(\boldsymbol{\varphi}, \varepsilon)| \le 1$, *and if* $\boldsymbol{\varphi}$ *verifies Eq. (G34), then it is given by Eq. (G35) up to* $2\pi\boldsymbol{\nu}$, $\boldsymbol{\nu} \in \mathcal{Z}^\ell$.

Observation. This is a "global theorem" involving an inversion on a large set, namely, $\mathcal{T}^\ell$. It can be improved to cover the case when $\mathbf{f}$ depends parametrically on some $\mathbf{A} \in \mathcal{R}^p$ so that $(\mathbf{A}, \boldsymbol{\varphi}) \to \mathbf{f}(\mathbf{A}, \boldsymbol{\varphi})$ is a C^∞ function on $\mathcal{R}^p \times \mathcal{T}^\ell$. Then if $\mathbf{f}$ verifies the assumptions of the corollary for each $\mathbf{A} \in V \subset \mathcal{R}^p$, one can check that $\mathbf{g} \in C^\infty(V \times \mathcal{T}^\ell)$, $\forall\, \varepsilon < \varepsilon_\ell$.

Proof. Let $0 \le \varepsilon < \frac{1}{4}$. The Jacobian matrices L, J of Eq. (G34) regarded as an implicit equation $\mathbf{F}(\boldsymbol{\varphi}, \boldsymbol{\varphi}') = \mathbf{0}$ in $\mathcal{R}^\ell \times \mathcal{R}^\ell$ near the solution $(\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_0 + \varepsilon\, \mathbf{f}(\boldsymbol{\varphi}_0))$, with $\boldsymbol{\varphi}_0$ given in $\mathcal{T}^\ell$, are

$$L_{ij} = \delta_{ij}, \qquad J_{ij} = L_{ij} + \varepsilon \frac{\partial f_i}{\partial \varphi_j}, \tag{G36}$$

and by assumption [see Eqs. (E2), (E3), and (E10)] and since $\varepsilon < \frac{1}{4}$:

$$(\ell - \frac{1}{4}) < |J| < (\ell + \frac{1}{4}), \qquad (\ell - \frac{1}{2}) < |J^{-1}| < (\ell + \frac{1}{2}), \tag{G37}$$

so that the constants B, C in Eq. (G31) can be chosen $B, C \ge (\ell - \frac{1}{2})^{-1}$. We now apply Corollary 3 to our equation near $(\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_0 + \varepsilon \mathbf{f}(\boldsymbol{\varphi}_0))$ by choosing $\delta = \sqrt{\varepsilon}$, say, and noting that from Eqs. (G21) and (G32), it follows that for ε small enough,

$$\frac{\delta}{4(\ell - \frac{1}{2})^2} \le \varrho, \qquad B\widetilde{\varrho} \le B\delta, \tag{G38}$$

Noting that $\delta \gg \varepsilon$, we see that Corollary 3 implies that as $\boldsymbol{\varphi}_0$ varies on $\mathcal{T}^\ell$, the point $\boldsymbol{\varphi}_0 + \varepsilon\, \mathbf{f}(\boldsymbol{\varphi}_0)$ also varies covering $\mathcal{T}^\ell$ if ε is small enough.

Furthermore, the map of Eq. (G34) is one to one, for ε very small, as a map of $\mathcal{T}^\ell$ onto itself. In fact, if $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2 \in \mathcal{T}^\ell$ and if the segment σ given by $t \to \boldsymbol{\varphi}_1 t + \boldsymbol{\varphi}_2(1-t)$, $t \in [0,1]$, is the shortest segment on $\mathcal{T}^\ell$ connecting $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$, we see that the points $\boldsymbol{\varphi}'_1 = \boldsymbol{\varphi}_1 + \varepsilon\, \mathbf{f}(\boldsymbol{\varphi}_1)$ and $\boldsymbol{\varphi}'_2 = \boldsymbol{\varphi}_2 + \varepsilon\, \mathbf{f}(\boldsymbol{\varphi}_2)$ can coincide $\mod 2\pi$ only if $\boldsymbol{\varphi}'_1 = \boldsymbol{\varphi}'_2$, if ε is small. [1]

Since $\mathbf{f}$ is periodic, the assumption that σ is the shortest path on $\mathcal{T}^\ell$ leading from $\boldsymbol{\varphi}_1$ to $\boldsymbol{\varphi}_2$ cannot be restrictive and, therefore, the map $\boldsymbol{\varphi} \to \boldsymbol{\varphi} + \varepsilon \mathbf{f}(\boldsymbol{\varphi})$ is one to one for $\varepsilon < 1$.

So the map of Eq. (G34) can be inverted on $\mathcal{T}^\ell$ and its inverse map $\boldsymbol{\varphi}' \to \mathbf{F}_\varepsilon(\boldsymbol{\varphi}')$ is C^∞ near every point if ε is small enough. Clearly, Eq. (G35) holds with $\mathbf{g}(\boldsymbol{\varphi}', \varepsilon) = -\mathbf{f}(\mathbf{F}_\varepsilon(\boldsymbol{\varphi}'))$ which also proves $|\mathbf{g}| < 1$. mbe

Concluding Remark
The above proofs do not really make use of the fact that $\mathbf{f}$ is of class C^∞. If $\mathbf{f}$ is only supposed to be of class $C^{(k)}$, $k \geq 1$, the ideas of the proofs still work, and the only difference will be that the inverse function $\boldsymbol{\varphi}$ will not turn out to be of class C^∞, of course, but only of class $C^{(k)}$. We use the above "$C^{(k)}$-version" of the implicit function theorems only in §5.7.

Exercise
In the context of Proposition 1, compute the second derivative of $f(\mathbf{x})$ in terms of f and of its first derivatives $\frac{\partial \varphi}{\partial \mathbf{x}}$ and in terms of f and of its first two derivatives. (*Answer:*

$$\frac{\partial^2 \varphi}{\partial x_j \partial x_i} = -\frac{\frac{\partial^2 f(\mathbf{x},\varphi)}{\partial x_i \partial x_j} + \frac{\partial^2 f(\mathbf{x},\varphi)}{\partial x_i \partial y} \cdot \frac{\partial \varphi}{\partial x_j}}{\frac{\partial f(\mathbf{x},\varphi)}{\partial y}} + \frac{\frac{\partial f(\mathbf{x},\varphi)}{\partial x_i}\big(\frac{\partial f(\mathbf{x},\varphi)}{\partial x_i \partial y} + \frac{\partial^2 f(\mathbf{x},\varphi)}{\partial y^2} \cdot \frac{\partial \varphi}{\partial x_i}\big)}{\big(\frac{\partial f(\mathbf{x},\varphi)}{\partial y}\big)^2} .)$$

6.8 H: The Ascoli-Arzelá Convergence Criterion

The following elegant proposition is famous.

1 Proposition. *Let Ω be a closed bounded set in $\mathcal{R}^d$. Let $(f_n)_{n=0}^\infty$ be a sequence of continuous functions defined on Ω such that:*
(i) The sequence $(f_n)_{n=0}^\infty$ is "equibounded", i.e., there exists M such that

$$||f_n|| = \max_{\boldsymbol{\xi} \in \Omega} |f_n(\boldsymbol{\xi})| \leq M \tag{H1}$$

(ii) the sequence $(f_n)_{n=0}^\infty$ is "equicontinuous", i.e., given $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$\sup_{n, |\boldsymbol{\xi} - \boldsymbol{\xi}'| < \delta_\varepsilon} |f_n(\boldsymbol{\xi}) - f_n(\boldsymbol{\xi}')| < \varepsilon. \tag{H2}$$

[1] In fact, $|\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2|$ cannot be too large ($\leq \pi$) if σ is the shortest segment joining $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$ on $\mathcal{T}^\ell$.

Then there is a subsequence $(f_{n_i})_{i=0}^{\infty}$ *such that the limit*

$$f(\boldsymbol{\xi}) = \lim_{i\to\infty} f_{n_i}(\boldsymbol{\xi}) \tag{H3}$$

exists, uniformly, $\forall\,\boldsymbol{\xi} \in \Omega$.

Observations.
(1) Hence f is continuous on Ω.
(2) The most interesting aspect of this theorem is the uniformly of the convergence.

Proof. Let $\Omega_0 \subset \Omega$ be a denumerable dense subset of Ω (to be concrete, think of the case when Ω is a square and Ω_0 is the set of its points with rational coordinates). We shall write $\Omega_0 = \{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots\}$.

By the equiboundedness condition, it will be possible to find a subsequence $(f_{n_i})_{i=0}^{\infty}$ of $(f_n)_{n=0}^{\infty}$ such that the limits

$$\lim_{i\to\infty} f_{n_i}(\boldsymbol{\xi}_j) \stackrel{def}{=} f(\boldsymbol{\xi}_j) \tag{H4}$$

exist. For instance, one can use the Cantor diagonal method; f is defined by the right-hand side of Eq. (H4).

Without loss of generality, we may and shall assume that the subsequence $(n_i)_{i=0}^{\omega}$ coincides with $(0, 1, 2\ldots)$, i.e., that the limits $\lim_{n\to\infty} f_n(\boldsymbol{\xi}_j)$ exist without passing to a subsequence. This will now be used to show that the function f defined on Ω_0, can be extended to Ω by showing that the limit $\lim_{n\to\infty} f_n(\boldsymbol{\xi})$ exists $\forall\,\boldsymbol{\xi} \in \Omega$.. In fact, we show that $(f_n(\boldsymbol{\xi}))_{n=0}^{\infty}$ is a Cauchy sequence for all $\boldsymbol{\xi} \in$
O.

Let $\boldsymbol{\xi} \in \Omega$. Given $\varepsilon > 0$ let $\widetilde{\boldsymbol{\xi}} \in \Omega - 0$ be such that $|\boldsymbol{\xi} - \widetilde{\boldsymbol{\xi}}| < \delta_\varepsilon$, see (ii); then, by Eq. (H2):

$$\begin{aligned}|f_n(\boldsymbol{\xi}) - f_m(\boldsymbol{\xi})| \le & |f_n(\boldsymbol{\xi}) - f_n(\widetilde{\boldsymbol{\xi}})| + |f_n(\widetilde{\boldsymbol{\xi}}) - f_m(\widetilde{\boldsymbol{\xi}})| \\ & + |f_m(\widetilde{\boldsymbol{\xi}}) - f_m(\boldsymbol{\xi})| \le 2\varepsilon + |f_n(\widetilde{\boldsymbol{\xi}}) - f_m(\widetilde{\boldsymbol{\xi}})| \xrightarrow[n,m\to\infty]{} 2\varepsilon\end{aligned} \tag{H5}$$

because $f(\widetilde{\boldsymbol{\xi}})_{n=0}^{\infty}$ is a Cauchy sequence. Hence, by the arbitrariness of e, we see that $(f_n(\boldsymbol{\xi}))_{n=0}^{\infty}$ is also a Cauchy sequence and we can define, $\forall\,\boldsymbol{\xi} \in \Omega$, $f(\boldsymbol{\xi}) = \lim_{n\to\infty} f_n(\boldsymbol{\xi})$.

If $\boldsymbol{\xi}, \boldsymbol{\eta} \in \Omega$, $|\boldsymbol{\xi} - \boldsymbol{\eta}| < \delta_\varepsilon$, t then follows from Eq. (H2) that

$$|f(\boldsymbol{\xi}) - f(\boldsymbol{\eta})| = \lim_{n\to\infty} |f_n(\boldsymbol{\xi}) - f_n(\boldsymbol{\eta})| \le \varepsilon \tag{H6}$$

It remains to show that the limit given by Eq. (H3) is uniform on Ω. Otherwise, we could find $\varepsilon > 0$, a sequence $n_i \xrightarrow[i\to\infty]{} \infty$ and points $x_i \in \Omega$ such that

$$|f_{n_i}(\mathbf{x}_i) - f(\mathbf{x}_i)| > \varepsilon, \qquad i = 1, 2, \ldots \tag{H7}$$

Assuming (no loss of generality) that $n_i = i$ i.e.,

$$|f_n(\mathbf{x}_n) - f(\mathbf{x}_n)| > \varepsilon, \qquad n = 1, 2, \dots \tag{H8}$$

This is impossible because there would be an accumulation point $\overline{\mathbf{x}} \in \Omega$ for the sequence x_n, $n = 1, 2 \dots$ and again we may assume, without loss of generality, that $\lim \mathbf{x}_n = \overline{\mathbf{x}}$. Then if $|\overline{\mathbf{x}} - \mathbf{x}_n| < \delta_{\frac{1}{4}\varepsilon}$, using Eqs. (H6) and (H2),

$$\begin{aligned} \varepsilon <&|f_n(\mathbf{x}_n) - f(\mathbf{x}_n)| \le |f_n(\mathbf{x}_n) - f(\overline{\mathbf{x}}_n)| + |f_n(\overline{\mathbf{x}}_n) - f(\overline{\mathbf{x}}_n)| \\ &+ |f(\overline{\mathbf{x}}) - f(\overline{\mathbf{x}}_n)| \le \frac{2\varepsilon}{4} + |f_n(\overline{\mathbf{x}}) + f_m(\overline{\mathbf{x}})| \xrightarrow[n\to\infty]{} \frac{1}{2}\varepsilon \end{aligned} \tag{H9}$$

which is a contradiction. mbe

2 Corollary. *Under the assumptions of Proposition 1, aside from that of boundedness (or of closure or both) for Ω, the same conclusions hold with the exception of the uniformity of the convergence of $f_{n_i}(\boldsymbol{\xi})$ to $f(\boldsymbol{\xi})$. Nevertheless, f is uniformly continuous on Ω.*

Proof. By inspection of the proof of Proposition 1.

Exercises

1. Let $(f_n)_{n=0}^{\infty}$ be a sequence of $C^{(1)}(\Omega)$ functions on a convex set which is the closure of its interior. If there is M such that $\sup_n \max_{\boldsymbol{\xi}\in\Omega} |\frac{\partial f_n(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}}| \le M$ then $(f_n)_{n=0}^{\infty}$ is an equicontinuous family on Ω. (*Hint:* Express the variation of f, as the integral of its derivative along a segment joining two points.)

2. Define $C^{(\varepsilon)}(\Omega)$, $\varepsilon \in (0, 1]$, to be the set of the functions such that

$$|f|_\varepsilon \stackrel{def}{=} \sup_{\mathbf{x}} |f(\mathbf{x})| + \sum_{\mathbf{x},\mathbf{y}} \frac{|f(\mathbf{x}) - f(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^e} < +\infty$$

Then any sequence$(f_n)_{n=0}^{\infty}$, $f_n \in C^{(\varepsilon)}(\Omega)$, such that $|f_n|_\varepsilon \le M < +\infty$, $\forall\, n$, is an equicontinuous equibounded sequence.

6.9 I: Fourier Series for Functions in $\overline{C}^{\infty}([0, L])$

Lemma 11, §4.5, p.266, will be proved here.

If $u \in \overline{C}^{\infty}([0, L])$, set

$$\begin{aligned} u^*(x) &= u(x), && x \in [0, L], \\ u^*(L + x) &= -\, u(L - x), && x \in [0, L] \end{aligned} \tag{I1}$$

and, by the assumption that the even derivatives of u in 0 and in L vanish, the function thus defined on $[0, 2L]$ is in $C^{\infty}([0, 2L])$ and is periodic, together with all its derivatives, with period $2L$. By the Fourier theorem, we set

$$\widehat{u}^*_h = \frac{1}{2L}\int_0^{2L} u^*(x)e^{-i\frac{2\pi h}{2L}x}\,dx \tag{I2}$$

for $h \in \mathcal{Z}$ and, $\forall\, 0 < h \in \mathcal{Z}_+$, remark that

$$\begin{aligned}\widehat{u}^*_h =&\frac{1}{2L}\int_0^L \big(u(x)e^{-i\frac{\pi h}{L}x} - u(L-x)e^{-i\frac{\pi h}{L}(L+x)}\big) -\, dx\\ =&\frac{1}{2L}\int_0^L u(x)\big(e^{-i\frac{\pi h}{L}x} - e^{i\frac{\pi h}{L}x}\big)\,dx\\ =&\frac{-i}{L}\int_0^L u(x)\sin\frac{\pi h x}{L}\,dx = \frac{-i}{2}\overline{u}(h) = -\widehat{u}^*_{-h},\end{aligned} \tag{I3}$$

having used the change of variables $x \to L - x$.. Therefore, for $x \in [0, 2L]$,

$$u^*(x) = \sum_{h=-\infty} +\infty \widehat{u}^*_h e^{i\frac{\pi h}{L}x} = \sum_{h=1} +\infty \overline{u}_h \sin\frac{\pi h}{L}x. \tag{I4}$$

Hence, for $x \in [0, L]$,

$$u(x) = \sum_{h=1}^{+\infty} \overline{u}_h \sin\frac{\pi h}{L}x., \tag{I5}$$

where $\overline{u}_h$ defined in Eq. (I3) coincides with Eq. (4.5.20). Equation (4.5.21) follows from Eq. (I3) and from the decay properties as $h \to \infty$ of the Fourier coefficients for C^∞-periodic functions. Equation (I5) gives Eq. (4.5.22). mbe

6.10 L: Proof of Eq. (5.6.20)

Let $(S_t^{(\alpha,\delta)}(w_1, w_2))_i \stackrel{def}{=} \sigma_i(t, \mathbf{w})$, $i = 1, 2$, $t \in [0, 1]$. Eqs. (5.6.17),(5.6.18) give

$$\begin{aligned}\sigma_1(t,\mathbf{w}) =&w_1 + \int_0^t \chi_\delta(\boldsymbol{\sigma}(\tau,\mathbf{w}))\,(\alpha\sigma_1(\tau,\mathbf{w}) + P(\boldsymbol{\sigma}(\tau,\mathbf{w})))\,d\tau,\\ \sigma_2(t,\mathbf{w}) =&e^{-\nu_0 t}w_2 + \int_0^t e^{-\nu_0(t-\tau)}\chi_\delta(\boldsymbol{\sigma}(\tau,\mathbf{w}))\,Q(\boldsymbol{\sigma}(\tau,\mathbf{w}))\,d\tau\end{aligned} \tag{L1}$$

.

Consider, for instance, $\frac{\partial\boldsymbol{\sigma}}{\partial w_1}$ and drop the $\mathbf{w}$ in the arguments of $\boldsymbol{\sigma}$, for simplicity. Note that

$$\begin{aligned}\frac{\partial\sigma_1}{\partial w_1} =&1+\int_0^t \Big\{\partial\chi_\delta(\boldsymbol{\sigma}(\tau))\cdot\frac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\,(\alpha\sigma_1(\tau)+P(\boldsymbol{\sigma}(\tau)))\\ &+\chi_\delta(\boldsymbol{\sigma}(\tau))\,(\alpha\frac{\partial\sigma_1(\tau)}{\partial w_1}+\partial P(\boldsymbol{\sigma}(\tau))\cdot\frac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1})\Big\}\,d\tau\\ \frac{\partial\sigma_2}{\partial w_1} =&\int_0^t e^{-\nu_0(t-\tau)}\Big\{\partial\chi_\delta(\boldsymbol{\sigma}(\tau))\cdot\frac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\,Q(\boldsymbol{\sigma}(\tau))\\ &+\chi_\delta(\boldsymbol{\sigma}(\tau))\,\partial Q(\boldsymbol{\sigma}(\tau))\cdot\frac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\Big\}\,d\tau\end{aligned} \tag{L2}$$

where ∂g denotes $(\frac{\partial g}{\partial w_1},\frac{\partial g}{\partial w_2})$ if g is a function of w_1, w_2 and possibly other variables. Hence, using Eq. (5.6.15) and the fact that P and Q have a second-order zero at the origin, we see that there are two constants p, q such that

$$\big|\tfrac{\partial\sigma_1(t)}{\partial w_1}-1\big|\le p\int_0^t\Big\{\tfrac{1}{\delta}\big|\tfrac{|\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\big|(|\alpha|\delta+\delta^2)+|\alpha|\big|\tfrac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\big|+|\delta|\big|\tfrac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\big|\Big\}d\tau \tag{L3}$$

(since $|\boldsymbol{\sigma}(\tau)|\le\delta\sqrt{2}$) and

$$\big|\frac{\partial\sigma_1(t)}{\partial w_2}\big|\le q\int_0^t\Big\{\frac{1}{\delta}\big|\frac{|\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\big|\delta^2+|\delta|\big|\frac{\partial\boldsymbol{\sigma}(\tau)}{\partial w_1}\big|\Big\} \tag{L4}$$

Therefore, adding and subtracting 1 appropriately:

$$\begin{aligned}\big|\frac{\partial\sigma_1(t)}{\partial w_1}-1\big| \le&2p(|\alpha|+\delta)t+2p(|\alpha|+\delta)\\ &\cdot\int_0^t\Big\{\big|\frac{|\partial\sigma_1(\tau)}{\partial w_1}-1\big|+\big|\frac{|\partial\sigma_1(\tau)}{\partial w_1}\big|\Big\}\,d\tau\\ \big|\frac{\partial\sigma_2(t)}{\partial w_1}\big| \le&2q+2q\delta\int_0^t\Big\{\big|\frac{|\partial\sigma_1(\tau)}{\partial w_1}-1\big|+\big|\frac{\partial\sigma_1(\tau)}{\partial w_1}\big|\Big\}\,d\tau\end{aligned} \tag{L5}$$

Setting $y(t)=|\frac{\partial\sigma_1(\tau)}{\partial w_1}|+|\frac{\partial\sigma_1(\tau)}{\partial w_1}|$, the preceding inequalities, added up, imply

$$y(t)\le 2(p+q)(|\alpha|+\delta)t+2(p+q)(|\alpha|+s)\int_0^t y(\tau)\,d\tau \tag{L6}$$

and $y(0)=0$. The above integral inequality implies $y(t)\le\overline{y}(t)$, $\forall\, t\ge 0$, where

$$\overline{y}(t)\le 2(p+q)(|\alpha|+\delta)t+2(p+q)(|\alpha|+\delta)\int_0^t \overline{y}(\tau)\,d\tau \tag{L7}$$

and $y(0)=0$ (see Problems 8 and 9, §2.5). Hence,

$$\overline{y}(t)=(e^{2(p+q)(|\alpha|+\delta)t}-1)\le Mt(|\alpha|+\delta) \tag{L8}$$

for $0\le t\le 1$, $|\alpha|\le 1$, $\delta\le 1$ (and M could be $2(p+q)e^{4(p+q)}$).

An identical argument could be given for $t\in(-1,0)$.

6.11 M: Proof of Eq. (5.6.63)

Let

$$(x, \pi_t(x)) = S_t^{(\alpha,\delta)}(x_0, \pi(x_0)),\ (x, \pi_{t'}(x)) = S_{t'}^{(\alpha,\delta)}(x'_0, \pi(x'_0)) \tag{$M1$}$$

Then from Eq. (5.6.33), it follows that

$$\begin{aligned} &|\pi_t(x) - \pi_{t'}(x)| \le |e^{-\nu_0 t}\pi(x_0) - e^{-\nu_0 t'}\pi(x'_0)| + \Big| \int_0^t d\tau\, e^{-\nu_0(t-\tau)} \\ &Z_\delta(S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha) - \int_0^{t'} d\tau\, e^{-\nu_0(t'-\tau)} Z_\delta(S_\tau^{(\alpha,\delta)}(x'_0, \pi(x'_0)), \alpha)\Big|. \end{aligned} \tag{$M2$}$$

Using Eqs. (5.6.25), (5.6.24), and (5.6.49) and supposing $0 \le t' < t \le t_+$, the right-hand side of Eq. (M2) is

$$\begin{aligned} \le& |e^{-\nu_0 t} - e^{-\nu_0 t'}||\pi(x_0)| + e^{-\nu_0 t'}|\pi(x_0) - \pi(x'_0)| \\ &+ M\delta^2|t - t'| + \int_0^{t'} |e^{-\nu_0(t-\tau)} - e^{-\nu_0(t'-\tau)}|\, M\delta^2\, d\tau \\ &+ \int_0^{t'} e^{-\nu_0(t'-\tau)}\, |Z_\delta(S_\tau^{(\alpha,\delta)}(x_0, \pi(x_0)), \alpha) - Z_\delta(S_\tau^{(\alpha,\delta)}(x'_0, \pi(x'_0)), \alpha)| d\tau \\ \le& \delta\nu_0|t - t'| + |\pi(x_0) - \pi(x'_0)| + (1 + \nu_0 t)M\delta^2|t - t'| \\ &+ 2Mt\delta(1 + M(a_+ + \delta)t)(|x_0 - x'_0| + |\pi(x_0) - \pi(x'_0)|) \\ \le& (\delta\nu_0 + (1 + \nu_0 t)M\delta^2)|t - t'| + \big\{\big(c\sqrt{\delta} + 2Mt\delta(1 + M(a_+ + \delta)t) \\ &+ 2Mt\delta(1 + M(a_+ + \delta)t)C\sqrt{\delta}\big)\big\}|x_0 - x'_0|. \end{aligned} \tag{M3}$$

To estimate $|x_0 - x'_0|$ proceed as in subsection 5.6.G, p.421, using the expressions analogous to Eq. (5.6.58):

$$\begin{aligned} x_0 =& x - \int_0^t d\tau X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi_t(x)), \alpha), \\ x'_0 =& x - \int_0^{t'} d\tau X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi_{t'}(x)), \alpha), \end{aligned} \tag{$M4$}$$

By Eqs. (5.6.25), (5.6.23), and (5.6.20),

$$\begin{aligned} &|x_0 - x'_0| \le M|t - t'|(a_+\delta + \delta^2) + \int_0^{t'} d\tau \\ &\cdot |X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi_t(x)), \alpha) - X_\delta(S_{-\tau}^{(\alpha,\delta)}(x, \pi_{t'}(x)), \alpha)| \\ &\le M(a_+\delta + \delta^2)|t - t'| + 2M(a_+ + \delta)t(1 + M(a_+ + \delta)t)|\pi_t(x) - \pi_{t'}(x)|. \end{aligned} \tag{M5}$$

The restrictions imposed on a, t_0, by the second of Eqs. (5.6.41) imply (recall that $C, \delta < 1$)

$$\begin{aligned}\theta =&(C\sqrt{\delta} + 2tM\delta(1 + M(a_+ + \delta)t)(1 + C\sqrt{\delta}))\\ &\cdot 2M(a_+ + \delta)\, t\, (1 + M(a_+ + \delta)t)\\ \le&(1 + \frac{1}{10}(1 + \frac{1}{20}))(1 + 1)\frac{1}{10}(1 + \frac{1}{10}) < \frac{1}{2}\end{aligned} \tag{M6}$$

By combining the last of Eqs. (M5) with the last of Eqs. (M3), it follows that

$$(1 - \theta)|\pi_t(x) - \pi_{t'}(x)| \le (\delta\nu_0 + (1 + \nu_0 t)M\delta^2)|t - t'|, \tag{M7}$$

so that, since $\theta < \frac{1}{2}$:

$$|\pi_t(x) - \pi_{t'}(x)| \le 2(\delta\nu_0 + (1 + \nu_0 t_+)M\delta^2)|t - t'| \tag{M8}$$

$\forall\, t, t' \in [0, t_+]$; hence, by Eq. (5.6.51), for all $t, t' \in \mathcal{R}_+$, $t' \le t$, $|t - t'| < t_+$.

6.12 N: Analytic Implicit Functions

The proofs of Propositions 20 and 21, §5.11, are based on the following idea. Let F be a holomorphic function of a single complex variable $z \in \Omega \subset \mathcal{C}$. Assume that its complex derivative, denoted by a prime in this section, $F'(z)$, does not vanish in Ω.

It is a consequence of the theory of power series that, as z' varies in a small vicinity of $z'_0 = F(z_0)$ and z varies close to z_0, the equation $z' = F(z)$ can be uniquely solved for z by a function I defined in a neighborhood U of z_0 and holomorphic in U:

$$F(I(z')) \equiv z' \tag{N1}$$

for all z' in U, and

$$I(F(z)) \equiv z \tag{N2}$$

for all z in a suitable neighborhood of z_0. The function I has Taylor coefficients in z'_0 which can be computed via a simple algorithm from those of F in z_0.

The function F will be invertible on the whole $F(\Omega)$ if and only if $F(z) \ne F(z')$ whenever $z \ne z'$. In this case the inverse function I will be holomorphic on $F(\Omega)$ and it will be the unique inverse of F defined on $F(\Omega)$.

A simple criterion implying that $F(z) \ne F(z')$ for $z \ne z'$ is the following. Suppose that for every pair $z, z' \in \Omega$ there is a smooth curve $\Lambda(z, z') \subset \Omega$ with length $|\Lambda(z, z')|$ bounded by

$$\Lambda(z, z') < \beta(\Omega)|z - z'|, \tag{N3}$$

where $\beta(\Omega)$ is a suitable constant. Then F will be a one to one map between Ω and $F(\Omega)$ if

$$\sigma = \beta(\Omega) \sup_{z\in\Omega} |F'(z) - 1| < 1 \tag{N4}$$

In fact (N4) implies

$$\begin{aligned} &|F(z) - F(z')| \equiv |\int_{\Lambda(z,z')} F'(\zeta)d\zeta \equiv |\int_{\Lambda(z,z')} d\zeta \\ &+ \int_{\Lambda(z,z')} (F'(\zeta) - 1)d\zeta| \geq |z - z'| - \sigma|z - z'| = (1-\sigma)|z - z'|. \end{aligned} \tag{N5}$$

Proposition 20 can be proved by using the above remarks. First consider the inversion problem for the equation

$$\varphi' = \varphi + g(\varphi) \mod 2\pi \tag{N6}$$

with $\varphi \in \mathcal{T}^\ell$ and g holomorphic on $C(\xi)$. Let $\overline{g}$ be the holomorphic extension of g to $C(\xi)$. Eq. (N6) can be written

$$z' = z\, e^{i\overline{g}(z)} \equiv F(z), \qquad z \in \mathcal{T}^1 \tag{N7}$$

Let $\delta \in (0,1)$, $\delta < \frac{1}{2}\xi$ (say); we regard (N7) as an equation for $z \in C(\xi - \delta)$, i.e., $\Omega = C(\xi-\delta)$ in the language of the above discussion.

Between any two points $z, z' \in C(\xi)$ draw a line $\Lambda(z,z')$ contained in $C(\xi)$ with length $\leq 2\pi|z - z'|$: i.e. $\beta(C(\xi-\delta))$, see Eq. (N3), can be taken $= 2\pi$. Hence Eq. (N7) can be inverted in $C(\xi-\delta)$ under the condition

$$2\pi \sup_{z\in C(\xi-\delta)} |e^{i\overline{g}(z)} - 1 + i\overline{g}'(z)e^{i\overline{g}(z)}| < 1 \tag{N8}$$

which also ensures that $F'(z) \neq 0$ because

$$F'(z) \equiv 1 + (e^{i\overline{g}(z)} - 1) + \overline{g}'(z)e^{i\overline{g}(z)}. \tag{N9}$$

The supremum in inequality (N8) is bounded dimensionally, as in (5.11.18):

$$2\pi\big((e^{|\overline{g}|_\xi} - 1) + e^{|g|_\xi}e^{\xi}\delta^{-1}\big) < 2\pi e^{2\xi}e^{|g|_\xi}|g|_\xi\delta^{-1} \tag{N10}$$

By the above analysis a function $I(z')$ on $F(C(\xi-\delta))$ can be defined with

$$F(I(z')) = z', \qquad \forall\, z' \in F(C(\xi-\delta)), \qquad \text{provided} \tag{N11}$$

$$4\pi e^{2\xi}e^{|g|_\xi}|g|_\xi\delta^{-1} < 1 \tag{N12}$$

The form of F,

$$F(z) = ze^{i\overline{g}(z)}, \tag{N13}$$

implies that

$$F(C(\xi-\delta)) \subset C(\xi-\delta-|g|_\xi) \tag{N14}$$

because the F-image of $\partial C(\xi-\delta)$ consists of two lines outside $C(\xi-\delta-\||g|_\xi)$ and the boundary of $F(C(\xi-\delta))$ is $F(\partial C(\xi-\delta))$. The latter property follows from general properties of holomorphic functions but it can also be seen directly in our case as follows. If $z'_0 \in F(C(\xi-\delta))$, there is a sequence $z_n \in C(\xi-\delta)$ such that

$$z'_0 = \lim n \to \infty F(z_n) \tag{N15}$$

and, without loss of generality, we may suppose that the sequence z_n converges to a limit z_0. If $z_0 \in \partial C(\xi-\delta)$, then $z'_0 \in F(C(\xi-\delta))$; if $z_0 \in C(\xi-\delta)$, then the local invertibility of F implies that z_0 is interior to $F(C(\xi-\delta))$ which is impossible.

Therefore if Eq. (N12) holds the function I inverse to F is holomorphic at least in $C(\xi-2\delta)$, because Eq. (N12) implies $|g|_\xi < \delta$. Assuming the validity of the inequality in Eq. (N12), set

$$\overline{\Delta}(z') = -\overline{g}(I(z')), \qquad z' \in C(\xi-2\delta). \tag{N16}$$

This defines a holomorphic function on $C(\xi-2\delta)$ such that

$$|\overline{\Delta}|_{\xi-2\delta} < |g|_\xi, \qquad \text{and} \tag{N17}$$

$$I(z') = z'\, e^{i\overline{\Delta}(z')}. \tag{N18}$$

As z varies on the unit circle, the point $z' = F(z)$ also varies on the unit circle so that $\overline{\Delta}$ is real on the F-image of the unit circle: since $|g|_\xi < \delta$ and F is given by Eq. (N13) it follows (by a continuity argument) that as z varies on the unit circle z' varies covering the entire unit circle. This means that $\overline{\Delta}$ is real on $\mathcal{T}^1$ and it becomes possible to define

$$\Delta(\varphi) \stackrel{def}{=} \overline{\Delta}(e^{i\varphi}) \tag{N19}$$

and Δ is analytic and real on $\mathcal{T}^1$.

Since δ is arbitrary in $(0, \frac{1}{2}\xi)$, replacing 2δ by δ the theorem is proved, in the case considered, under the condition

$$8\pi e^\xi e^{|g|_\xi} |g|_\xi \delta^{-1} < 1 \tag{N20}$$

Next we study the inversion problem for the equation

$$\varphi' = \varphi + g(\mathbf{A}, \varphi), \tag{N21}$$

where g is holomorphic on $C(\varrho, \xi; \mathbf{A}_0)$. We write Eq. (N21) as

$$z' = z\, e^{i\overline{g}(\mathbf{A},z)}, \tag{N22}$$

Repeat, $\forall\, \mathbf{A} \in \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, the above argument, once more keeping in all the formulae an explicit $\mathbf{A}$ dependence which will, however, play no role whatsoever. So Eq. (N22) will be invertible in the form

$$z = z'\, e^{i\overline{g}(\mathbf{A},z)} \tag{N23}$$

with $\overline{\Delta}$ holomorphic on $C(\varrho, \xi - 2\delta; \mathbf{A}_0)$ if Eq. (N20) holds with $|g|_\xi$ replaced by $|g|_{\varrho,\xi}$. The function $\overline{\Delta}$ will also turn out to be real for $\mathbf{A} \in \mathcal{S}_\varrho(\mathbf{A}_0)$, i.e. for $\mathbf{A}$ real and for $|z'| = 1$.

The same conclusions hold if g is defined and holomorphic on a more general set of the form $W \times \mathcal{T}^1$ with $W \subset \mathcal{C}^\ell$ open. Eq. (N22) is inverted by Eq. (N23) if Eq. (N20) holds with $|g|_\xi$ replaced by the supremum of g in $W \times C(x)$.

With these remarks in mind, the proof of Proposition 20 can be concluded. Consider the case contemplated in Proposition 20:

$$\boldsymbol{\varphi}' = \boldsymbol{\varphi} + \mathbf{g}(\mathbf{A}, \boldsymbol{\varphi}) \tag{N24}$$

with $\mathbf{g}$ extending to a holomorphic function on $C(\varrho, \xi; \mathbf{A}_0)$. Write the system of Eq. (N24) as

$$z_k' = z_k e^{i g_k(\mathbf{A},\mathbf{z})}, \qquad k = 1, \ldots, p, \tag{N25}$$

and consider the first equation for z_1:

$$z_1' = z_1 e^{i g_1(\mathbf{A}, z_1, \ldots, z_p)}. \tag{N26}$$

If Eq. (N20) holds with $|g|_{\varrho,\xi}$ replaced by $|\mathbf{g}|_{\varrho,\xi}$, we can invert Eq. (N26) as

$$z_1 = z_1' e^{i\widetilde{\Delta}_1(\mathbf{A}, z_1', \ldots, z_p)} \tag{N27}$$

with $\widetilde{\Delta}$, holomorphic for $\mathbf{A} \in \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, $z_1' \in C(\xi - \delta)$, and $z_k \in C(\xi)$ for all $k = 2, \ldots, p$. Also, $|g|_{\varrho,\xi} < \delta$. Furthermore, Eq. (N27) inverts Eq. (N26) on the same set $\mathbf{A} \in \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0) \times C(\xi - \delta) \times C(\xi)^{\ell-1}$, and

$$|\widetilde{\Delta}_1| < |g_1|_{\varrho,\xi} \le |\mathbf{g}|_{\varrho,\xi} \tag{N28}$$

where $|\widetilde{\Delta}_1|$ denotes the supremum of $\widetilde{\Delta}_1$ on its domain of definition. Finally, $\widetilde{\Delta}_1$, is real if $\mathbf{A} \in \mathcal{S}_\varrho(\mathbf{A}_0)$, $|z_1'| = |z_2| = \ldots = |z_k = 1$.

Now substitute Eq. (N27) into the Eq. (N25) for $k = 2, \ldots, p$, and set

$$g_k^{(1)}(\mathbf{A}, z_1', z_2 \ldots, z_p) = g_k(\mathbf{A}, z_1' e^{i\widetilde{\Delta}_1(\mathbf{A}, z_1', z_2 \ldots, z_p)}, z_2, \ldots, z_p) \tag{N29}$$

which are defined and holomorphic for $\mathbf{A} \in \widehat{\mathcal{S}}_\varrho(\mathbf{A}_0)$, $z_1' \in C(\xi - \delta)$, and $z_k \in C(\xi)$ and, of course, the supremum of $|g_k^{(1)}|$ on its domain of definition can be estimated as

$$\sup |g_k^{(1)}| \le |g_k^{(1)}|_{\varrho,\xi} \le |\mathbf{g}|_{\varrho,\xi} \tag{N30}$$

Hence, we can take as parameters $\mathbf{A}, z_1', z_2 \ldots, z_p$) and solve the equation

$$z_2' = z_2\, e^{ig_2^{(1)}(\mathbf{A},z_1',z_2,\ldots,z_p)} \tag{N31}$$

for z_2 as before, etc. After p steps, we will have inverted the full system, in the desired form, on the set $C(\varrho, \xi - \delta; \mathbf{A}_0)$ under the sole condition

$$8\pi e^{2\xi} e^{|\mathbf{g}|_{\varrho,\xi}} |\mathbf{g}|_{\varrho,\xi} |\delta^{-1} < 1, \tag{N32}$$

Which, if $\delta < 1$, and, hence, $|g|_{\varrho,\xi}| < \gamma$ can be put into the form of Eq. (5.11.19) with $\gamma < 2^8$. With some care, one could find smaller values for γ.

mbe

In the same way, one can prove the implicit function theorem mentioned in Proposition 21. Since this is a "local theorem", the proof is actually slightly easier than the above.

6.13 O: Finite-Difference Method

Consider $\mathbf{f} \in C^\infty(\mathcal{R}^d)$ and the equation

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{O1}$$

To estimate $\mathbf{x}(\tau)$, given $\tau > 0$, let $\eta = \frac{1}{N}\tau, N \in \mathcal{Z}_+$, and define inductively

$$\begin{aligned} \mathbf{x}_0 =& \mathbf{x}(0), \\ \mathbf{x}_n =& \mathbf{x}_{n-1} + \eta\, \mathbf{f}(\mathbf{x}_{n-1}), \qquad n = 1, 2, \ldots, N. \end{aligned} \tag{O2}$$

Let

$$C = \sup_{\mathbf{x}\in\Omega} \sum_{i=1}^{d} |f_i(\mathbf{x})|, \qquad L = \sup_{\mathbf{x}\in\Omega} \sum_{i,j=1}^{d} \Big|\frac{\partial f_i(\mathbf{x})}{\partial x_j}\Big|, \tag{O3}$$

where $\Omega \subset \mathcal{R}^d$ is some convex region where one can a priori guarantee that $\mathbf{x}(t)$, $\forall\, t \in [0, \tau]$, and $\mathbf{x}_n$, $\forall\, n = 0, 1, \ldots, N$, will fall ($\Omega$ has to be found in each case: out of despair one could always take $\Omega = \mathcal{R}^d$). Then

$$|\mathbf{x}_N - \mathbf{x}(\tau)| \le \frac{C\tau}{2N}\,(e^{L\tau} - 1) \tag{O4}$$

This formula gives an a priori estimate of the error that would be committed if one iteratively solved Eq. (O1) with the method of Eq. (O2) ("finite-difference algorithm"). It can be used in many of the exercises proposed in this book, where the use of a computer is suggested.

The proof of Eq. (O4) is a simple consequence of the considerations and proofs given in §2.2-§2.4.

Proof. Let $\mathbf{d}_k \stackrel{def}{=} \mathbf{x}_k - \mathbf{x}(k\eta),\ k = 0, 1, \ldots, N$. One finds

$$\begin{aligned} \mathbf{d}_k =& \mathbf{x}_{k-1} + \eta \mathbf{f}(\mathbf{x}_{k-1}) - \mathbf{x}((k-1)\eta) - \int_0^\eta \mathbf{f}(\mathbf{x}((k-1)\eta + \theta))\, d\theta \\ =& \mathbf{d}_{k-1} - \int_0^\eta \big(\mathbf{f}(\mathbf{x}((k-1)\eta + \theta)) - \mathbf{f}(\mathbf{x}_{k-1})\big)\, d\theta. \end{aligned} \tag{O5}$$

Hence, applying Taylor's formula and adding and subtracting suitable terms:

$$\begin{aligned} |\mathbf{d}_k| \le& |\mathbf{d}_{k-1}| + L \int_0^\eta |\mathbf{x}((k-1)\eta + \theta) - \mathbf{x}_{k-1}| d\theta \\ \le& |\mathbf{d}_{k-1}| + L \int_0^\eta (|\mathbf{x}((k-1)\eta + \theta) - \mathbf{x}((k-1)\eta)| + |\mathbf{d}_{k-1}) d\theta \\ \le& |\mathbf{d}_{k-1}| + L\eta |\mathbf{d}_{k-1}| + LC \int_0^\eta \theta d\theta. \end{aligned} \tag{O6}$$

where in the last inequality, the derivative of $\mathbf{x}$ has been bounded by recalling that $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and $|\mathbf{f}(\mathbf{x})| \le C$. If $\Omega \ne \mathcal{R}^d$ Taylor's formula can still be applied by the convexity assumption on Ω (by the proofs of appendix A). Then

$$|\mathbf{d}_k| \le (1 + L\eta)|\mathbf{d}_{k-1}| + \frac{LC}{2}\eta^2 \tag{O7}$$

which, by iteration, yields (since $\mathbf{d}_0 = \mathbf{0}$)

$$|\mathbf{d}_k| \le \frac{LC}{2}\eta^2 \sum_{j=0}^{k-1} (1 + L\eta)^j = \frac{C\eta}{2}[(1 + L\eta)^k - 1] \tag{O8}$$

which for $k = N$, recalling that $\eta = \frac{\tau}{N}$, becomes

$$|\mathbf{x}_N - \mathbf{x}(\tau)| \le \frac{C\tau}{2N}\big[(1 + \frac{L\tau}{N})^N - 1\big] \le \frac{C\tau}{2N}(e^{L\tau} - 1) \tag{0.9}$$

The approximation is therefore of order $O(N^{-1})$ at fixed τ. Since the relation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, by differentiating $n - 1$ times with respect to t, yields expressions for the first n derivatives it is possible to obtain "higher order approximations", $O(N^{-n})$, by natural modifications of the above algorithm.

It is also possible to achieve higher order approximations avoiding the (often lengthy) calculations of the higher order derivatives and using only $\mathbf{f}(\mathbf{x})$ (of course evaluated at several points): the most common algorithm is the Runge-Kutta algorithm. Its fourth order version is used in producing the graphs of §4.8 in the programs attached to this book.

6.14 P: Astronomical Data

(1) Gravitational constant $k = 6.67 \times 10^{-8} cm^3/g(sec)^2$.
(2) Radius of the Sun: $R_s = 6.96 \times 10^5 Km$.
Mass of the Sun: $M_s = 1.99 \times 10^3 3g$.
Density of the Sun: $p_s = 1.41 g/cm^3$. (3) Elements of the Planets' Orbits.

Planet	*Semiaxis* *u.a.*	*Semiaxis* 10^6 *km*	*SiderealPeriod* *Days*	*Eccentricity*	*Eclipticincl*	*Long.* *Asc.node*	*Long.* *Perigee*
Mercury	0.387099	57.91	87.969	0.206625	$7^o0'13''.8$	$47^o44'66"$	$76^o40'32"$
Venus	0.723332	108.21	224.700	0.006793	32339.3	761411	1305120
Earth	1.000000	149.60	365.257	0.016729		1020441	
Mars	1.52369	227.94	686.980	0.093357	1510.0	491025	3355819
Jupiter	5.2028	778.34	4332.587	0.048417	11821.2	995655	133133
Saturn	9.540	1427.2	10759.21	0.055720	22926.1	1131337	920439
Uranus	19.18	2869.3	30685.	0.0471	04622.0	734336	16951
Neptune	30.07	4498.5	60188.	0.0087	14628.1	1311351	4410
Pluto	39.44	5900.	90700.	0.247	170824	1093802	22330

For the year 1950. From [5]

(3) Elements of the Planets' Orbits.

Planet	*Radius* *km*	*Radius* */Earth*	*Mass* */Earth*	*Mass* 10^{27} *g*	*Density* g/cm^3	*Grav.* *accel.*	*Escape* *Km/s*	*Period* *sideral*	*Equator's* *inclin.*
Mercury	2437	0.382	0.055	0.330	5.5	372	4.3	$58d.65$	$7o$
Venus	6050	0.950	0.816	4.87	5.2	887	10.4	$243d.2**$	$3o24'$
Earth	6378	1.000	1.000	5.98	5.5	981	11.2	$23h56'4''.1$	$23o27'$
Mars	3394	0.531	0.107	0.64	3.9	376	5.0	$24h37'22''.6$	$24o56'$
Jupiter	71400	11.2	318.	1900.	1.3	2500	61.	$9h50'.5$	$3o07'$
Saturn	60400	9.5	95.1	568.	0.7	1100	36.	$10h14'$	$26o45'$
Uranus	24800	3.9	14.6	87.	1.6	950	22.	$1Oh49'**$	$82o$
Neptune	25050	3.9	17.2	103.	1.7	1150	24.	$15h.81$	$29o$
Pluto	¡2900	0.45	0.9*	5.5*	–	–	–	$6d.4$	–

* Approximate ** Retrograde
From [5].

(5) Satellites of the Planets.

Planet	*Satellite*	*Av.distance* $10^3 km$	*PeriodSid.* *days*	*PeriodSyn.* *days*	*Inclin.*	*Eccen.*	*Radius* *km*	*Mass* *(Pl./Sat)*	*Mass* $1O^{24} g$
Earth	*Moon*	384.4	27.321661	$29d12h44'02''.8$	*5lE*	0.0549	1738	81.3	73.4
Mars	1.*Phobos*	9.4	0.318910	073926.65	1.8*P*	0.019	14		
	2.*Deimos*	43.5	1.262441	1062115.68	1.4*P*	0.003	8		
Jupiter	1.*Io*	421.8	1.769138	1182835.95	*OP*	*Small*	1660	24000	79
	2.*Europa*	671.4	3.551181	3131753.74	*OP*	*and*	1440	39800	7.8
	3.*Ganymede*	1071.	7.154553	7035935.86	*OP*	*variab.*	2470	12400	153
	4.*Callisto*	1884.	16.689018	16180506.92	*OP*		2340	21000	90
	5.*Amalthea*	181.	0.498179	115727.6	*OP*	0.003	80		
	6.	11500.	250.62	260.0	28.5*B*	0.155	60		
	7.	11750.	259.8	276.10	28.0*B*	0.207	20		
	8.	23500.	738.9	631.05	*R*33*B*		0.38	20	
	9.	23700.	755.	626	*R*24*B*		0.25	11	
	10.	11750.	260.	276	28.3*B*	0.140	10		
	11.	22500.	696.	599	*R*16.6	0.207	12		
	12.	21000.	625.	546	*R*	0.13	10		
Saturn	1.*Mimas*	185.7	0.942422	223712.4	1.5*P*	0.0196	260	15000000	0.038
	2.*Encelado*	238.2	1.370218	1085321.9	*O.OP*	0.0045	300	8000000	0.07
	3.*Tethys*	294.8	1.887802	1211854.8	1.1*P*	0.0000	600	870000	0.65
	4.*Dione*	377.7	2.736916	2174209.7	*O.OP*	0.0021	650	555000	1.03
	5.*Rhea*	527.5	4.517503	4122756.2	0.3*P*	0.0009	900	250000	2.3
	6.*Titan*	1223.	15.945452	15231525	0.3*P*	0.0289	2500	4150	137
	7.*Hyperion*	1484.	21.276665	21073906	0.6*P*	0.110	200	5000000	0.11
	8.*Iapetus*	3563.	79.33082	79220456	14.7*P*	0.029	600	100000	5
	9.*Phoebe*	12950.	550.45	53616	*R*30*P*	0.166	150		
	10.*Themis*	157.5	0.749			300			
Uranus	1.**A***riel*	191.8	2.52038	2122940	*OP*	0.007	300		
	2.*Umbriel*	267.3	4.14418	4032825	*OP*	0.008	200		
	3.*Titania*	438.7	8.70588	81700	*OP*	0.023	500		
	4.*Oberon*	586.6	13.46326	13111536	*OP*	0.010	400		
	5.*Miranda*	130.1	1.414						
Neptune	1.*Triton*	353.6	5.87683	5210327	*R*20*P*	0.000	2000	700	150
	2.*Nereid*	6000?.	500.			0.7	150	3000000	0.05

P. on the plane of the planet's equator
B. on the plane of the planet's orbit
R. retrograde rotation
From [5].

6.15 Q: Gauss Method for Planetary Orbits of an Orbit through Three Observations

This appendix contains a series of guided problems on the two body central motion which is taken from the Gauss' treatise on the motion of heavenly bodies gravitating about the Sun in conic sections (1804).

1. *(Earth motions)* the Earth is assumed spherical and its rotation axis has a conical precession motion around the axis $\mathcal{N}$ *celestial north*, perpendicular to the Earth orbital plane ε, *ecliptic*. The two rotations take place at angular velocities, respectively, ω_D and ω_p. The velocities are called the *diurnal* rotation and the *precessional* rotation. The second is very slow for the following qualitative reasons which could be made quantitative at least as far as the orders of magnitude are concerned and even as far as the actual theoretical computation of the first order corrections.

Show that if the Earth was really a perfect sphere then one would expect that the Earth axis would stay fixed in orientation (*Hint:* in a frame of reference with center at the Earth center and axes fixed with the fixed stars the moment of the forces exercised by the Sun and by the Moon would vanish by symmetry and so would the moment of the inertial forces. Hence the motion would be that of a sphere with fixed center and no external forces: i.e. the axis would be fixed and no precession would be present).

Show also that if the Earth had cylindrical symmetry around its axis and one still neglected the forces exercised by the Sun and the Moon, then one would expect it to have a uniform rotation around its axis which in turn would rotate at constant angular velocity around a fixed axis (oriented as the angular momentum), keeping a constant angle with it.

2. *(further considerations on the Earth spin motion)* the following heuristic considerations are useful to keep in mind, even though strictly speaking, they are not specifically part of the problem of the orbit determination but rather pertain to the general problem of fixing the reference frames.

Since the Earth angular velocity and angular momenta can be taken as essentially parallel this movement would simply cause the inclination of the Earth axis as well as the intersection between the ecliptic and the plane orthogonal to the Earth axis to have a small motion around their average values: it could not be responsible for the precession motion. At best it could account for a small motion of the rotation axis around its average position. The precession is therefore caused by the action of the forces due to the Sun and Moon and to the non spherical symmetry of the Earth, and to the non circularity of the Earth and Moon orbits (causing further variations of the forces exercised by the Sun and Moon).

If the forces due to the Sun and to the Moon and to the inertial forces were constant in time in the frame of reference with center at the Earth and x-axis pointing at the Sun (which they are not because the distance and relative positions of the bodies change periodically in time, to a first approximation) then the Earth motion would be that of a top subject to a constant torque

moment trying to put the Earth equatorial plane on the ecliptic plane, where the Sun and Moon can be thought to be (again to a first approximation). Hence, as it follows from the theory of the spinning top, the Earth axis rotates around the $\mathcal{N}$ axis (because the variations of the angular velocity have to rotate around the axis such that if the body was oriented parallel to it then the moment of the forces would vanish, which in our case is $\mathcal{N}$ since the Earth is compressed at the poles).

However in the theory of the top it emerges that the speed of rotation is not uniform: it follows in fact that it periodically changes with time. Hence the motion is only to a first approximation uniform and the actual motion consists of the above rotation-precession plus a *nutation* motion which causes the precession speed to be altered and the inclination to oscillate quasi periodically around a mean value. All the above corrections can be given explicit theoretical values by using the theory of perturbation of integrable motions in the assumption that the motions of the Sun and of the Moon are essentially known and given by the Kepler's laws: this is called the *principal correction.*

Again this may not be satisfactory and one could introduce further refinements. We do not enter here into the details of such calculations and we summarize the above discussion by saying that one can compile, on theoretical grounds tables which allow to determine as a function of time the positions of the Earth axis and that of the intersection of the Earth equatorial plane and the ecliptic plane. The data given below are deduced from such tables, called the *Astronomical Ephemeris tables.*

3. *(zenithal frame)* if O is an astronomical observatory the local system of coordinates will have the origin in O and z-axis pointing upwards vertically (along a plomb line), i.e. towards the *zenith* Z. The x-axis will be the *horizon* axis $\varOmega$, determined by the tangent to the Earth in the plane μ of the z-axis and the terrestrial axis or *north axis*; the orientation of the $\varOmega$-axis will be towards south. The plane μ containing the zenith axis and the north axis will be called the *meridian plane.* Draw a graphical representation of the above frame.

4. *(equatorial frame)* in this frame one takes the origin to be the Earth center T the z-axis to be the axis N of the Earth's rotation oriented towards north. The plane ZN cuts the plane orthogonal to the N axis (called the *equatorial plane*) along a line called the *equator* line which is taken to be the x-axis of the equatorial frame.

Thus the equatorial frame and the zenithal frame are fixed relative to each other. Check that the angle δ_O between the equator and the zenith is what is commonly called the *latitude* of the observatory and draw a graphical representation of the zenithal and equatorial frames.

5. *(geocentric frame)* in this frame the origin is the Earth center T but the xy plane is the ecliptic plane (see 1)). The z axis points to the celestial north $\mathcal{N}$ and the x-axis will be parallel to the intersection between the equatorial

plane and the ecliptic plane ε. The orientation of the axis is towards the *Aries* constellation, the Γ point, (in fact this axis goes roughly through the Spring and Autumn noon positions of the Sun, i.e. through *Aries* and *???*).

When the Sun crosses this axis one has the *equinox*, respectively the Spring or Autumn equinoxes. The x axis is called the *equinox* axis and is denoted by Γ. The angle between the axes $\mathcal{N}$ and N axes is the *inclination* angle i_0 of the Earth axis. Therefore the Γ axis is not fixed in direction but rotates around the $\mathcal{N}$ axis with angular velocity ω_p.

Find a graphical representation for the equatorial and the geocentric frames.

6. *(heliocentric frame)* this is an inertial frame: its origin is at the center of mass of the solar system (which we confuse here with the center of the Sun for simplicity). The z axis is orthogonal to the ecliptic and parallel to the $\mathcal{N}$ axis previously introduced. Thus the xy plane is the ecliptic plane ε. The x axis will be parallel to the equinox axis Γ.

Knowing that the Earth motion on the ecliptic is a counterclockwise rotation, as seen standing up on the northern emisphere, check that when the Earth crosses the positive Γ-axis it is the Autumn equinox and that the N axis is obtained from the $\mathcal{N}$ axis by a clockwise rotation of an angle equal to the Earth *inclination* angle i_0 (*Hint:* because the Sun is in *Aries* in Spring and because Winter comes after Autumn).

Find a graphical representation of the heliocentric and of the geocentric frames and mark the point where the Earth would be at the Autumn or Spring equinox and the Γ point.

7. (precession and nutation) since the Γ point moves because of the precession one fixes the x axis of the heliocentric and geocentric systems to be the above axes in the positions in which they were at a given time called the *epoch* E of the time measurements, which is taken as the origin of the time. At any other time t the position of the Γ axis will form an angle $\omega_p(t-E)$ with the x axis of the heliocentric system. To distinguish between the two lines one calls the actual intersection between the equator and the ecliptic the *apparent* equinox line, denoted Γ_{app}.

The Γ-point and inclination i_0 in fact change in time also because of the nutation and we assume for the purposes of this illustration of Gauss method that this change can be desumed from the astronomical ephemeris tables to be equivalent to replacing λ_p by $\lambda_p + \lambda_n$ and ∞_o by ∞_n.

8. *(observations)* by *observation* of a celestial body one means the recording of the time t at which it crosses the meridian plane and of the angle δ above the horizon on which it is seen at the moment of the crossing. Sometimes one records instead of δ the angle δ_E at which it is seen. Of course the relation between the two data is simply: $\delta_E = \delta + \delta_O - \pi/2$. The angles δ and δ_E are the *heights* above the horizon or above the equator. Show that one expects to

have to make three observations to determine the orbit of the planet (*Hint:* the system has three degrees of freedom, i.e. six parameters).

9. *(apparent positions)* call $\mathbf{R}_T$ the vector leading from S to T, $\mathbf{R}_O$ the vector from T to O and $\mathbf{B}_a$ the unit vector pointing in the apparent position of the body. In the zenithal frame it is $\mathbf{B}_a = (\cos\delta, 0, \sin\delta)$. Introduce the matrices:

$$V_1(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} V_2(\alpha) = \begin{pmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} V_1(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (Q1)$$

and the vectors

$$\mathbf{n}_1 = (1,0,0), \mathbf{n}_2 = (0,1,0), \mathbf{n}_3 = (0,0,1)$$

check that the matrices V_1, V_2, V_3 rotate the whole world counterclockwise by an angle α around the axes $1, 2, 3$.

We call λ_T the angle between T and the Γ_{app}-line; the angle of inclination of the Earth axis will be i_0. Because of the mentioned the precession and nutation the angle i between the N axis and the $\mathcal{N}$ axis is somewhat different from i_0. Also the longitude angle between T and the fixed Γ line is $\lambda_T + \lambda_p + \lambda_n$ (see 7) above). Let R be the Earth radius.

Show that:

$$\begin{aligned} \mathbf{R}_T &= D_T V_3(\lambda_T + \lambda_p + \lambda_n)\mathbf{n}_1 \\ \mathbf{R}_O &= R V_3(\lambda_p + \lambda_n) V_1(-i) V_3(\lambda_O) V_2(\frac{\pi}{2} - \delta_O)\mathbf{n}_3 \\ \mathbf{B}A &= V_3(\lambda_p + \lambda_n) V_1(-i) V_3(\frac{\pi}{2} - \delta_O + \frac{\pi}{2} - \delta)\mathbf{n}_3 \end{aligned} \quad (Q2)$$

Setting $\mathbf{A} = \mathbf{R}_T + \mathbf{R}_O$, $\mathbf{X}_a = \mathbf{A} + \varrho\mathbf{B}_a$ where ϱ is the distance between the the heavenly body C and the observatory O, we see that the vector $\mathbf{A}$ consists of two terms of different order of magnitude (because $R/D_T << 1$): the first is the heliocentric *place*of the Earth and the second is the *parallax* correction. The vector $\mathbf{B}_a$ is the apparent heliocentric place of the heavenly body and ϱ is of course unknown.

10. *(fixed stars aberration)* this is a further correction that bears this name because it has to be considered even when one observes a fixed star. It is due to the finiteness of the light speed c. By the composition law of the classical velocity we see that if $c\mathbf{B}$ is the light velocity in the heliocentric frame and $c'\mathbf{B}'$ is the velocity in the zenith frame and if $\mathbf{v}$ is the velocity of the observatory, then:

$$c\mathbf{B} = c'\mathbf{B}_a + \mathbf{v} \quad (Q3)$$

Show that if one neglects corrections of order $(v/c)^2$ then one can write:

$$\mathbf{B} = \mathbf{B}_a|\mathbf{B} - \frac{\mathbf{v}}{c}| + \frac{\mathbf{v}}{c} \simeq \mathbf{B}_a(1 - \frac{\mathbf{B}_a \cdot \mathbf{v}}{c}) + \frac{\mathbf{v}}{c} \quad (Q4)$$

There is no need to use the relativistic velocity composition law as it leads to corrections of order $O((v/c)^2)$ which have anyway been neglected in deducing the last equation.

11. *(computation of fixed stars aberrations)* if ω_T is the diurnal Earth period show that the velocity of the observatory can be written:

$$\mathbf{v}_O = \omega_T R \cos\delta_O \, V_3(\lambda_p + \lambda_n) V_1(-i) V_3(\lambda_O) \mathbf{n}_2 \tag{Q5}$$

Show also that $\mathbf{v}_T$ can be computed from the fact that the Earth motion is Keplerian in terms of the vector $\mathbf{R}_T$, which in the heliocentric frame has polar coordinates D_T, $\lambda_T + \lambda_n + \lambda_p$, by:

$$\mathbf{v}_T = \dot D_T \frac{\mathbf{R}_T}{D_T} + D_T \dot\theta \frac{\mathbf{n}_3 \wedge \mathbf{R}_T}{D_T} \tag{Q6}$$

Using Eqs. (4.10.11),(4.10.12),(4.10.18), i.e. the fact that the areas constant is $A = 2\pi R_g R_m / T$ where T is the Earth revolution period and R_g, R_m are the great axis of the Earth orbit and the minor axis, show that:

$$\begin{aligned} \dot D =& \pm \frac{2\pi R_g R_m}{T} \Big((\frac{1}{R_-} - \frac{1}{D_T})(\frac{1}{D_T} - \frac{1}{R_+}) \Big)^{1/2} \\ D_T \dot\theta =& \frac{2\pi R_g R_m}{T} \frac{1}{D_T} \end{aligned} \tag{Q7}$$

where R_+, R_- denote the perihelion and aphelion distances in the Earth orbit, i.e. $R_\pm = R_g(1 \pm e)$ if e is the Earth orbit eccentricity.

Setting $\bar D = \frac{D_t}{R_g}$ we find after some algebra:

$$\mathbf{v}_T = \pm \frac{2\pi(1-e^2)^{1/2}}{T} \Big((\frac{1}{1-e} - \frac{1}{\bar D})(\frac{1}{\bar D} - \frac{1}{1+e}) \Big)^{1/2} \frac{\mathbf{R}_T}{\bar D} + \frac{1}{\bar D} \frac{\mathbf{n}_3 \wedge \mathbf{R}_T}{D_T} \tag{Q8}$$

The sign to choose in (Q.8) is $-$ for observations between roughly the summer solstice and the winter solstice and $+$ in the other period (as in this epoch the perihelion is early in January a few days after the winter solstice).

The calculation of the fixed stars aberrations needs not be computed if one has astronomical tables containing in some form its value, for the observatory of interest.

12. *(time aberrations)* If $\mathbf{A} + \varrho \mathbf{B}$ is the heliocentric position calculated as above as a function of the unknown distance ϱ between the Earth and the heavenly body, and a s observed at the time t one has to think that in fact it provides us with the position really occupied by the heavenly body at the time $t - \varrho/c$, since the speed of light is finite.

Furthermore sometimes the astronomical tables give the geocentric position of the Sun rather than the heliocentric position of the Earth: in this case some obvious changes have to be made to the above formulae and one has to add the further correction on the time of the observation obtained by reading

in the tables the data relative to the times $t+t_s$ if t_s is the time necessary to the light to travel from the Sun to the Earth, i.e. 500^s or $\approx 8^m$. In practice this means that one has to change the Earth longitude λ_T into $\lambda_T+\tilde{\lambda}$ if $\tilde{\lambda}$ is the arc described by the Earth in the time t_s: this would be constant if the Earth had a circular orbit, but it varies in a way that can be desumed from the tables around a mean value of 20.25^s, with oscillations between -0.34^s (at the perihelion) and $+0.34^s$ (at the aphelion).

13. *(planetary aberrations)* the Earth ecliptic plane (as one should by now suspect) is in fact also not fixed in space, mainly because of the perturbations caused by the Jupiter attraction, and the Earth is not exactly on the ecliptic plane, mainly because of the Moon (in fact, it is the center of mass of the Earth-Moon system which is really moving and defining the ecliptic plane): hence the Sun has an *apparent latitude* $-\beta$: this small quantity is directly measurable (for the main Moon contribution) or is accessible to theoretical analysis and can be found in the tables. One can take it into account (neglecting terms of $O(\beta^2)$) simply correcting the expression for the vector $\mathbf{A}$ by adding to it a vector $+\beta D_T \nu_3$.

14. *(summary of the heliocentric coordinates calculations)* a heavenly body C observed on the meridian with height above the equator δ_E at a time t is the sum of two vectors $\mathbf{A}$ and $\varrho\mathbf{B}$ whose Cartesian components components in a heliocentric system can be computed, via the astronomical tables which provide the orbital data for the Earth. the aberrations etc. In terms of the symbols introduced in the previous problems one finds:

$$\begin{aligned}\mathbf{A} =& D_T V_3(\lambda_T+\lambda_p+\lambda_n)\mathbf{n}_1 + R V_3(\lambda_p+\lambda_n)\\ &\cdot V_1(-i)V_3(\lambda_O)V_2(-\delta_O)\mathbf{n}_1 + \beta D_T \mathbf{n}_3\\ \mathbf{B}_a =& V_3(\lambda_p+\lambda_n)V_1(-i)V_3(\lambda_O)V_2(-\delta_e)\mathbf{n}_1\\ \mathbf{B} =& \mathbf{B}_a\left(1-\frac{\mathbf{v}_T\cdot\mathbf{B}_a}{c}\right)+\frac{\mathbf{v}_T+\mathbf{v}_O}{c}\end{aligned} \tag{Q9}$$

The ϱ coordinate is not measurable directly: it will be our main problem to show that it can be computed from the data.

Compute $\mathbf{A},\mathbf{B}$ from the following table providing the data of the asteroid Juno (observed at Greenwich on October 5, 17, 27 1904; the data (taken from the book of Gauss) are referred to the epoch E =1 January 1805):

t	δ_E	λ_T	λ_p	$i-i_0$
$5^d10^h51^m6^s$	$-6^o40^p8^s$	$12^o28^p53.72^s$	11.87^s	59.48^s
$17^d9^h58^m10^s$	$-8^o47^p25^s$	$24^o20^p21.54^s$	10.23^s	59.26^s
$27^d9^h16^m41^s$	$-10^o2^p28^s$	$34^o16^p52.21^s$	8.86^s	59.06^s
λ_O	D_T	β	λ_n	

$357^o10^p22.35^s$	0.9988899	0.49^s	-15.43^s
$355^o43^p45.30^s$	0.9953968	-0.79^s	-15.41^s
$355^o11^p10.95^s$	0.9928340	0.15^s	-15.60^s

and:

$R = 4.1683397 \times 10^{-5}$	$e = 0.016729$	$\delta_O = 51^o28^p39^s$
$i_0 = 23^o27^p$	$c = 2.0039603 \times 10^{-3} au/s$	$T_D = 23^h56^m4.1^s$
$R_g = 1.496 \times 10^8\, km$	$T = 3.15582048 \times 10^7 s$	

where D_T and R are in astronomical units and T_D is the period of rotation of the Earth, (use a computer to write a program producing the Cartesian components of $\mathbf{A}$ and $\mathbf{B}$).

15. *(planarity condition)* let $\mathbf{A}_i, \mathbf{B}_i, \mathbf{X}_i,\ i = 1, 2, 3$ be the vectors describing aberration free heliocentric coordinates of a heavenly body gravitating around the Sun according to the Keplerian laws. Then $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ are on the same plane. Show that this implies:

$$(\mathbf{X}_1 \wedge \mathbf{X}_2 \cdot \mathbf{k})\mathbf{X}_3 + (\mathbf{X}_2 \wedge \mathbf{X}_3 \cdot \mathbf{k})\mathbf{X}_1 + (\mathbf{X}_3 \wedge \mathbf{X}_1 \cdot \mathbf{k})\mathbf{X}_2 = \mathbf{0} \qquad (Q10)$$

where $\mathbf{k}$ denotes the unit vector orthogonal to the plane $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ (*Hint:* remark that there exist α, β such that $\mathbf{X}_3 = \alpha\mathbf{X}_1 + \beta\mathbf{X}_2$ and substitute in (Q.14)).

If one introduces the oriented areas $n_{pq}/2$ of the triangles $S_{pq}, p, q = 1, 2, 3$ formed by joining S, p, q, show that (Q.14) becomes:

$$n_{12}\mathbf{X}_3 + n_{23}\mathbf{X}_1 - n_{13}\mathbf{X}_2 = \mathbf{0} \qquad (Q11)$$

because of the geometrical meaning of $\mathbf{X}_p \wedge \mathbf{X}_q \cdot \mathbf{k}$.

16. *(distance and area relations)* show that (Q.11) implies:

$$\begin{aligned}
a\varrho_1 &= -\left(\mathbf{B}_2 \wedge \mathbf{B}_3 \cdot \mathbf{A}_1\right) + \frac{n_{13}}{n_{23}}(\mathbf{B}_2 \wedge \mathbf{B}_3 \cdot \mathbf{A}_2) - \frac{n_{12}}{n_{23}}(\mathbf{B}_2 \wedge \mathbf{B}_3 \cdot \mathbf{A}_3) \\
a\varrho_2 &= -\frac{n_{23}}{n_{13}}(\mathbf{B}_1 \wedge \mathbf{B}_3 \cdot \mathbf{A}_1) + (\mathbf{B}_1 \wedge \mathbf{B}_3 \cdot \mathbf{A}_2) - \frac{n_{12}}{n_{13}}(\mathbf{B}_1 \wedge \mathbf{B}_3 \cdot \mathbf{A}_3) \qquad (Q12) \\
a\varrho_3 &= -\frac{n_{23}}{n_{12}}(\mathbf{B}_1 \wedge \mathbf{B}_2 \cdot \mathbf{A}_1) + \frac{n_{13}}{n_{12}}(\mathbf{B}_1 \wedge \mathbf{B}_2 \cdot \mathbf{A}_2) - (\mathbf{B}_1 \wedge \mathbf{B}_2 \cdot \mathbf{A}_3)
\end{aligned}$$

where $a = (\mathbf{B}_1 \wedge \mathbf{B}_2) \cdot \mathbf{B}_3$).

17. *(other distance areas relations)* an alternative set of relations. which will be useful is found my multiplying the first of the (Q.11) vectorially by $\mathbf{B}_3$ (thus eliminating the explicit dependence on ϱ_3 and then scalarly by $\mathbf{B}_1 \wedge \mathbf{B}_3$ (so that in the same sense one eliminates ϱ_3). Show that in this way one finds:

$$(\mathbf{B}_1 \wedge \mathbf{B}_3)^2 \varrho_1 = -\frac{n_{12}}{n_{23}}(\mathbf{A}_3 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) - (\mathbf{A}_1 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) +$$
$$+ \frac{n_{13}}{n_{23}}(\mathbf{A}_2 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) + \frac{n_{13}}{n_{23}}\varrho_2(\mathbf{B}_2 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) +$$
$$(\mathbf{B}_1 \wedge \mathbf{B}_3)^2 \varrho_3 = (\mathbf{A}_3 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) + \frac{n_{13}}{n_{12}}(\mathbf{A}_1 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) -$$
$$- \frac{n_{13}}{n_{12}}(\mathbf{A}_2 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) - \frac{n_{13}}{n_{12}}\varrho_2(\mathbf{B}_2 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) + \tag{Q13}$$

18. *(computation of relevant constants)* with reference to the above two problems write a program for the computation of the following constants:

$$\begin{aligned}
a =&(\mathbf{B}_1 \wedge \mathbf{B}_2) \cdot \mathbf{B}_3 \qquad b = (\mathbf{B}_1 \wedge \mathbf{B}_3) \cdot \mathbf{A}_2 \\
c =&-(\mathbf{B}_1 \wedge \mathbf{B}_3) \cdot \mathbf{A}_1 \quad d = -(\mathbf{B}_1 \wedge \mathbf{B}_3) \cdot \mathbf{A}_3 \\
\gamma_0 =&(\mathbf{B}_1 \wedge \mathbf{B}_3)^2 \\
\gamma_1 =&-(\mathbf{A}_3 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \qquad \gamma_2 = -(\mathbf{A}_1 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \\
\gamma_3 =&(\mathbf{A}_2 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \qquad \gamma_4 = (\mathbf{B}_2 \wedge \mathbf{B}_3) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \\
\gamma_5 =&(\mathbf{A}_3 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \qquad \gamma_6 = (\mathbf{A}_1 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \\
\gamma_7 =&-(\mathbf{A}_2 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3) \qquad \gamma_8 = -(\mathbf{B}_2 \wedge \mathbf{B}_1) \cdot (\mathbf{B}_1 \wedge \mathbf{B}_3)
\end{aligned} \tag{Q14}$$

and show that the second of (Q.12) and (Q.13) become:

$$\begin{aligned}
a\varrho_2 =&b + c\frac{n_{23}}{n_{13}} + d\frac{n_{12}}{n_{13}} \\
\gamma_0\varrho_1 =&\gamma_1\frac{n_{12}}{n_{23}} + \gamma_2 + \frac{n_{13}}{n_{23}}(\gamma_3 + \varrho_2\gamma_4) \\
\gamma_0\varrho_3 =&\gamma_5 + \frac{n_{23}}{n_{12}}\gamma_6 + \frac{n_{13}}{n_{12}}(\gamma_7 + \varrho_2\gamma_8)
\end{aligned} \tag{Q15}$$

19. *orders of magnitude* suppose that the angles between $\mathbf{B}_i, \mathbf{B}_j$ are small and so are the angles between $\mathbf{A}_i, \mathbf{A}_j$, let ε be their order of magnitude. Show that the coefficients in the preceding problem have the following orders of magnitude in terms of ε:

$$\begin{array}{llll}
a = O(\varepsilon^3) & b = O(\varepsilon) & c = O(\varepsilon) & d = O(\varepsilon) \\
\gamma_0 = O(\varepsilon^2) & & & \\
\gamma_1 = O(\varepsilon) & \gamma_2 = O(\varepsilon) & \gamma_3 = O(\varepsilon) & \gamma_4 = O(\varepsilon^2) \\
\gamma_5 = O(\varepsilon) & \gamma_6 = O(\varepsilon) & \gamma_7 = O(\varepsilon) & \gamma_8 = O(\varepsilon^2)
\end{array} \tag{Q16}$$

(*Hint:* to see that $a = O(\varepsilon^3)$ note that the volume of the parallelepiped generated by three vectors forming an angle $O(\varepsilon)$ between each other is in general of $O(\varepsilon^2)$; however if $\mathbf{A}_1 = \mathbf{A}_2 = \mathbf{A}_3$ it would be $a = 0$, because then the $\mathbf{B}$'s would be in the same plane. Hence the reason why the $\mathbf{B}$'s are not on the same plane is because the $\mathbf{A}$'s are not identical; but the $\mathbf{A}$ and $\mathbf{B}$ vary smoothly and the $\mathbf{A}$'s too form between each other angles of $O(\varepsilon)$...).

20. *(necessary accuracy)* if $t_i, i = 1, 2, 3$ are the observation times let $t_{pq} = t_q - t_p$ and show that the Kepler laws imply:

$$\frac{n_{pq}}{n_{rs}} = \frac{t_{pq}}{t_{rs}} + O(\varepsilon^2) \tag{Q17}$$

if ε has the meaning of the previous problem. (*Hint:* the third Kepler's law gives proportionality between the signed area of the elliptic sector swept by C in the time t_{pq} and the area of the sector differs from that of the triangles by $O(\varepsilon^2)$).

Show that if one neglects $O(\varepsilon^2)$ and n_{pq}/n_{rs} is replaced by t_{pq}/t_{rs} (which is directly accessible from the measurements) in (Q.17) then one makes an error on ϱ_2, for instance, of O(1)! hence we see that we have to find a better way to start an approximation.

21. *(how well should ϱ_2 be known)* show that if ϱ_2 were known to $O(\varepsilon)$ then the second and third of $(Q.15)$ would permit us to evaluate ϱ_1, ϱ_3 also to an error of $O(\varepsilon)$ even using the approximation in which one makes an error of order $O(\varepsilon^2)$ in the ratio's n_{pq}/n_{rs}, (for instance replacing it by t_{pq}/t_{rs}; (*Hint:* this follows immediately from the estimates in 18)).

Therefore one has to look for an approximation of ϱ_2 within $O(\varepsilon)$.

22. *(Gauss' lemma)* introduce the ratios z_{pq} between the double of the area of the elliptic sector swept by C between the times t_p, t_q and the quantities n_{pq} introduced in 19) above. As already remarked such ratios differ from 1 by $O(\varepsilon^2)$ and furthermore the ratios $z_{pq}n_{pq}/t_{pq}$ are constant in p, q.

Consider the first expression for ϱ_2 in (Q.15) and show that if one replaces it with :

$$a\varrho_2 = b + \frac{ct_{23} + dt_{12}}{t_{23} + t_{12}} \frac{n_{23} + n_{12}}{n_{13}} \tag{Q18}$$

one makes an error on ϱ_2 of order $o(\varepsilon)$, rather than $o(1)$ (as one could believe on first thought on the basis of an argument similar to the one suggested in 18), until one remarked that:

$$\frac{ct_{23} + dt_{12}}{t_{23} + t_{12}} - \frac{cn_{23} + dn_{12}}{n_{23} + n_{12}} = \frac{t_{12}t_{23}(c - d)(z_{12} - z_{23})}{(t_{23} + t_{12})(z_{12}t_{23} + z_{23}t_{12})} \tag{Q19}$$

and that the denominator has size $O(\varepsilon^2)$ while the numerator has size $O(\varepsilon^4)(c - d)$ and, furthermore, that although c, d have size of $O(\varepsilon)$ their difference has size $O(\varepsilon^2)$ because $c - d = -(\mathbf{B}_2 - \mathbf{B}_3) \cdot (\mathbf{A}_1 - \mathbf{A}_3)$.

22. *(Gauss ϱ_2 equation)* let $\kappa = 2\pi R_g^{3/2}/T$, where R_g is the great semiaxis of any major planet orbiting the Sun and T is the corresponding period: it follows from (4.10.7) that κ^2 is the product of the sun mass times the universal gravitational constant. It follows from the theory of the two body problem, as remarked by Gauss that the following basic relation between n_{pq}, z_{pq}, $r_q = |\mathbf{X}_q|$ and the angles at S of the triangles S_{pq} with vertices in the Sun and the C positions at the tines of the corresponding observations:

$$\frac{n_{23}+n_{12}}{n_{13}} = 1 + \frac{\kappa t_{12}\kappa t_{23}}{2z_{12}z_{23}r_1r_2r_3\cos f_{12}\cos f_{23}\cos f_{13}} \qquad (Q20))$$

A guide to the derivation of the above relation is provided in the following problem 23). Approximating z and the cosines with 1 and identifying r_1, r_2, r_3 show that one finds:

$$a\varrho_2 = b + \frac{ct_{23}+dt_{12}}{t_{23}+t_{12}}\big(1+\frac{\kappa t_{12}\kappa t_{13}}{2r_2^3}\big) \qquad (Q21)$$

which is an equation for the unknown ϱ_2, because: t_{pq} are known and $r_2 = |\mathbf{A}_2+\varrho_2\mathbf{B}_2| = (\mathbf{A}_2^2+\varrho_2^2+2\varrho_2\mathbf{A}_2\cdot\mathbf{B}_2)^{1/2}$. We neglect here the time aberrations which will be corrected only at the end.

Usually (Q.21) admits only one acceptable solution (show, however that it can be rationalized and becomes an equation of eight degree). Show that it determines ϱ_2 to $O(\varepsilon)$ (*Hint:* the z_{pq} and the cosines differ from 1 by $O(\varepsilon^2)$ and the r_q differ between each other by $O(\varepsilon)$; furthermore $c, d, \kappa t_{pq}$ have size of $O(\varepsilon)$ and $a = (\varepsilon^3)\ldots$). Write a computer program to solve the equation (Q.21) and find its positive solutions.

23. *(digression on the two body problem to prove (Q.20)* rewrite Eq. (4.10.6) in the form:

$$r = \frac{p}{1-e\cos\theta} \qquad (Q22)$$

where p is the *parameter* of the ellipse and e is its eccentricity. Then deduce that:

$$\begin{aligned} &p = \frac{2\varrho_+\varrho_-}{\varrho_+ + \varrho_-} \quad e = \frac{\varrho_+ - \varrho_-}{\varrho_+ + \varrho_-} \quad a = \frac{\varrho_+ + \varrho_-}{2} \quad b = \sqrt{\varrho_+\varrho_-} \\ &\varrho_\pm = a(1\pm e) \quad b = a\sqrt{1-e^2} \qquad p = \frac{b^2}{a} \end{aligned} \qquad (Q23)$$

where we are denoting: a = major semiaxis of the ellipse, b = minor semiaxis.

If we call $\theta_1, \theta_2, \theta_3$ the three angles that $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ form with respect to a given line drawn on their plane (*eg* with respect to the ascending node with the ecliptic plane oriented parallel to $\mathbf{n}_3 \wedge (\mathbf{X}_1 \wedge \mathbf{X}_3)$, and if g denotes the angle between the same reference line and the major semiaxis of the elliptic orbit of the heavenly body, then the *true anomalies* of the three positions will be $\beta_1 = \theta_1 - g, \beta_2 - g, \beta_3 - g$ and it will be:

$$\begin{aligned} pr^{-1} =&(1-e\cos\beta_1) \\ pr^{-2} =&(1-e\cos\beta_2) \\ pr^{-3} =&(1-e\cos\beta_3) \end{aligned} \qquad (Q24)$$

Note that with the notations of 22) it is

$$\beta_3-\beta_2=\theta_3-\theta_2=2f_{23}, \beta_3-\beta_1=\theta_3-\theta_1=2f_{13}, \beta_2-\beta_1=\theta_2-\theta_1=2f_{12}$$

and check that:

$$\frac{z_{rs} n_{rs}}{\kappa t_{rs}} = \sqrt{p} \qquad (Q25)$$

furthermore, since $n_{pq} = r_p r_q \sin 2f_{pq}$, it is:

$$p = \frac{\sin 2f_{23} + \sin 2f_{12} - \sin 2f_{13}}{r_1^{-1} \sin 2f_{23} + r_2^{-1} \sin 2f_{12} - r_3^{-1} \sin 2f_{13}} = \frac{4 \sin f_{23} \sin f_{13} \sin f_{12} r_1 r_2 r_3}{n_{23} + n_{12} - n_{13}} \qquad (Q26)$$

and deduce from (Q.26),(Q.25) that (Q.20) holds (*Hint:* check first the (Q.26) by remarking that, by the second Kepler law, the ratio between the area of the elliptic sector S_{rs} and the area of the ellipse, i.e. $z_{rs} n_{rs}/2\pi ab$ coincides with the ratio between the time needed to sweep the sector and the heavenly body period, i.e. $t_{rs}/2\pi a^{3/2}/\kappa$ by the (4.10.7); hence (Q.25) and (Q.24) immediately imply (Q.26)). Then to get $(Q.20)$ one combines (Q.25),(Q.26), getting:

$$\begin{aligned} p = \frac{z_{23} z_{12} n_{23} n_{12}}{\kappa t_{23} \kappa t_{12}} &= \frac{4 \sin f_{23} \sin f_{13} \sin f_{12} r_1 r_2 r_3}{n_{23} + n_{12} - n_{13}} = \\ &= \frac{\sin 2f_{23} \sin 2f_{12} \sin 2f_{13} r_1 r_2 r_3}{2(n_{23} + n_{12} - n_{13}) \cos f_{23} \cos f_{12} \cos f_{13}} \end{aligned} \qquad (Q27)$$

i.e. :

$$\begin{aligned} n_{23} + n_{12} - n_{13} &= \frac{\kappa t_{23} \kappa t_{12}}{z_{23} z_{12} n_{23} n_{12}} \frac{\sin 2f_{23} \sin 2f_{12} \sin 2f_{13} r_1^2 r_2^2 r_3^2}{\cos f_{23} \cos f_{12} \cos f_{13} r_1 r_2 r_3} = \\ &\equiv \frac{\kappa t_{23} \kappa t_{12} n_{13}}{2 z_{12} z_{13} \cos f_{23} \cos f_{12} \cos f_{13} r_1 r_2 r_3} \end{aligned} \qquad (Q28)$$

because $n_{pq} = r_p r_q \sin 2f_{pq}$).

24. *(summary of above)* the preceding problems permit us to compute a first approximation ϱ_i^0 to the distances ϱ_i up to errors of order $O(\varepsilon)$, if ε is an estimate of the size of the angles between the vectors $\mathbf{A}_i$ or the vectors $\mathbf{B}_i$. Hence we have a first approximation $\mathbf{X}_i^0$ for the vectors $\mathbf{X}_i$. It is useful to summarize the above procedure as follows.

The value of ϱ_2^0 is found by solving the equation:

$$a\varrho_2 = b + \frac{ct_{23} + dt_{12}}{t_{23} + t_{12}} \big(1 + \frac{\bar{Q}}{2r_2^3}\big) \qquad (Q29)$$

where $\bar{Q} = \kappa t_{23} \kappa t_{12}$, determining ϱ_2 to $O(\varepsilon)$, (see problem 22). Set also $\bar{P} = t_{13}/t_{23}$ and:

$$\bar{W} = \bar{P}(1 + \frac{\bar{Q}}{2r_2^3} - \frac{1}{\bar{P}}) \qquad (Q30)$$

then one realizes that $\bar{P}, \bar{W}$ are approximations to n_{13}/n_{23} and n_{12}/n_{23} to order $O(\varepsilon^2)$. In fact this has been seen in (Q.17) for $\bar{P}$, and $\bar{W}$ differs from n_{12}/n_{23} because, (see (Q.20)), $(1 + \bar{Q}/2r_2^3)$ is not $(n_{12} + n_{23})/n_{13} = 1 + \bar{Q}/(2r_1 r_2 r_3 z_{12} z_{23} \cos f_{12} \cos f_{13} \cos f_{23}$ nor $1/\bar{P}$ is n_{23}/n_{13}, but the latter two

quantities differ from the preceding ones respectively to $O(\varepsilon)\bar{Q}$ and ε^2) (see the hint to problem 22) and the equation (Q.17)) and $\bar{Q}$ is of $O(\varepsilon^2)$)

Hence by problem 20) it is possible to find ϱ_1, ϱ_3 within $O(\varepsilon)$ by using the last two relations in (Q.15):

$$\begin{aligned}\gamma_0\varrho_1 =&\gamma_1\bar{W} + \gamma_2 + (\gamma_3 + \gamma_4\varrho_2)\bar{P} \\ \gamma_0\varrho_3 =&\gamma_5 + \gamma_0\bar{W}^{-1} + (\gamma_7 + \gamma_8\varrho_2)\bar{P}\bar{W}^{-1}\end{aligned} \tag{Q31}$$

25. *(elliptic elements)* at this point we can compute five parameters ($i^0, \lambda^0, g^0, e^0, p^0$) needed to determine the Keplerian orbit of the heavenly body using the information that is an ellipse with focus in the Sun S passing through the three points determined by the vectors $\mathbf{X}_i^0$: this will be the first approximation to the *elements* of the celestial body. We omit the superscript 0 in what follows to simplify the notations. The five *elements* are:

- i inclination of the orbit plane over the ecliptic
- λ longitude of the ascending node between the orbit and the ecliptic
- g angle between the orbit major axis oriented towards the aphelion and the ascending node
- e orbit eccentricity
- p ellipse parameter

Denoting $\mathbf{m}$ the versor of $\mathbf{X}_1^0 \wedge \mathbf{X}_3^0$ and with $\mathbf{m}'$ that of $\mathbf{n}_3 \wedge \mathbf{m}$ it is clear that $\mathbf{m}$ is normal to the orbit plane (by construction the $\mathbf{X}_i^0$ have been constructed to verify approximately the (Q.11), hence to be almost on the same plane) while $\mathbf{m}'$ is the ascending node between the orbit plane and the ecliptic.

Let $\theta_1, \theta_2, \theta_3$ be the angles formed, respectively, by $\mathbf{X}_1^0, \mathbf{X}_2^0, \mathbf{X}_3^0$ with the ascending node $\mathbf{m}'$: they are the angular polar coordinates of the three approximate positions in orbit, measured on the orbit plane with respect to the ascending node. Let $r_i = |\mathbf{X}_i|$ and check the following relations:

$$\cos i = \mathbf{n}_3 \cdot \mathbf{m} \qquad \cos\lambda = \mathbf{n}_1 \cdot \mathbf{m}' \qquad \sin\lambda = \mathbf{n}_2 \cdot \mathbf{m}' \tag{Q32}$$

and:

$$\begin{aligned}\tan(g) =& -\frac{r_1^{-1}(\cos\theta_2 - \cos\theta_3) + r_2^{-1}(\cos\theta_3 - \cos\theta_1) + r_3^{-1}(\cos\theta_1 - \cos\theta_2)}{r_1^{-1}(\sin\theta_2 - \sin\theta_3) + r_2^{-1}(\sin\theta_3 - \sin\theta_1) + r_3^{-1}(\sin\theta_1 - \sin\theta_2)} \\ e =& \frac{r_1^{-1} - r_3^{-1}}{r_1^{-1}\cos(\theta_3 - g) - r_3^{-1}\cos(\theta_1 - g)} \\ p =& \frac{\cos(\theta_3 - g) - \cos(\theta_1 - g)}{r_1^{-1}\cos(\theta_3 - g) - r_3^{-1}\cos(\theta_1 - g)}\end{aligned} \tag{Q33}$$

where the ambiguity on the g, defined up to π, is to be solved by imposing that the eccentricity e be positive; alternatively one can express p via (Q.26), *etc.* (*Hint:* use the (Q.24) in the form:

$$\begin{aligned} r_1^{-1} =& p^{-1} - ep^{-1}\cos(\theta_1 - g) \\ r_2^{-1} =& p^{-1} - ep^{-1}\cos(\theta_2 - g) \\ r_3^{-1} =& p^{-1} - ep^{-1}\cos(\theta_3 - g) \end{aligned} \qquad (Q34)$$

The tangent of g is found by multiplying the (Q.34) respectively by $(\cos\theta_2 - \cos\theta_3)$, $\cos\theta_3 - \cos\theta_1)$ and $\cos\theta_1 - \cos\theta_2)$ and adding the resulting equations side by side: the term with p^{-1} disappears and, developing the $\cos(\theta_i - g)$ via the addition formulae the terms with $\cos g$ also simplify. Repeat the scheme by multiplying by $(\sin\theta_2 - \sin\theta_3)$, *etc*: this time the terms with p^{-1} and $\sin g$ disappear; dividing the two relations thus obtained one finds the first of the (Q.33). Once g is known one finds p^{-1} and ep^{-1} from the first and third of the (Q.24), for instance, and one gets the last two of (Q.33). One could find other essentially equivalent expressions: for instance p can be determined also via the (Q26); (they would be really identical if the there had been no approximations)).

Write a computer program for the calculation of the five elements defined above.

26. *(consistency problems)* express in terms of $\mathbf{X}_i^0$, $i = 1, 2, 3$ the value that the ratios z_{pq}^0 between the areas of the elliptic sectors S_{pq} and the corresponding triangles take in the ellipse constructed in problem 24), assuming that the celestial body moves on it according to the Kepler laws and following the hints given below.

Let a, b, p be the major, minor axes of the ellipse and the parameter; if r_q, θ_q are defined as in problem 24), introduce the quantities β_q, ξ_q, l_q as follows:

$$\beta_q = \theta_q - g \qquad r_q = p(1 - e\cos\beta_q)^{-1} = a(1 + e\cos\xi_q) \qquad (Q35)$$

The above quantities are called *true anomaly*, it is the polar coordinate of $\mathbf{X}_q^0$ with respect to the major semiaxis), *eccentric anomaly* and *mean anomaly* of $\mathbf{X}_q^0$. Check that:

$$z_{pq}^o = \frac{ab(l_q - l_p)}{r_p r_q \sin(\theta_q - \theta_p)} \qquad (Q36)$$

(*Hint:* the average anomaly l is independently defined as the product of $2\pi/T$, T being the orbital period of the celestial body, times the time elapsed since the celestial body passed its aphelion: this notion, naturally arising in the theory of the central motions was defined in (4.9.31), where it was denoted φ_1 but setting the origin at the perihelion (hence the two definitions differ by π). On the basis of this definition one has, therefore:

$$\frac{dl}{dt} = \frac{2\pi}{T} \qquad l = 0 \quad \text{if} \quad \beta = 0 \qquad (Q37)$$

The (Q.36) is an immediate consequence of this property of the average anomaly which makes it proportional to the time elapsed since the passage through the aphelion. In the Keplerian motion the latter time is proportional

to the area of the elliptic sector swept by the celestial body, hence the area swept between T_p and t_q is to the area of the ellipse as the variation of the average anomaly is to 2π: i.e. the area swept is $\pi ab(l_q - l_p)/2\pi$. Since, obviously, the area of the triangle corresponding to the elliptic sector S_{pq} is $t_p r_q(\sin\theta_q - \theta_p)/2$ the (Q36) follows.

The true problem is therefore to check the (Q.35), once the average anomaly is defined via (Q.37). Recalling (4.10.11),(4.10.12), one finds, using (Q.37):

$$\begin{aligned} &\frac{dl}{dt} = \frac{2\pi}{T} \qquad \frac{d\beta}{dt} = \frac{A}{r^2} \\ &\frac{dr}{dt} = \pm A\sqrt{(\varrho_-^{-1} - \varrho^{-1})(\varrho^{-1} - \varrho_+^{-1})} \end{aligned} \tag{Q38}$$

where ϱ_-, ϱ_+ denote the distances of the perihelion and of the aphelion.

From the (4.10.16), and (4.10.18) one deduces the following relations between the areas constant A, the period T, *etc.*:

$$\begin{aligned} &\frac{1}{2}A = \frac{\pi ab}{T} \qquad a = \frac{\varrho_+ + \varrho_-}{2} \qquad b = \sqrt{\varrho_+\varrho_-} \\ &e = \frac{\varrho_+ - \varrho_-}{\varrho_+ + \varrho_-} \quad \varrho_\pm = a(1 \pm e) \quad p = \frac{b^2}{a} = a(1 - e^2) \end{aligned} \tag{Q39}$$

most of which have already been remarked in (Q.23). Hence the (Q.38) can be recast in the form:

$$\begin{aligned} &\frac{dl}{dt} = \frac{2\pi}{T} \qquad \frac{d\beta‘}{dt} = \frac{2\pi}{T}\frac{ab}{r^2} \\ &\frac{dr}{dt} = \pm\frac{2\pi ab}{T}\sqrt{\frac{(\varrho_+ - r)(r - \varrho_-)}{\varrho_+\varrho_- r^2}} = \pm\frac{2\pi a}{T}\frac{\sqrt{a^2e^2 - (r-a)^2}}{r} \end{aligned} \tag{Q40}$$

which imply, by dividing between each other conveniently the above relations:

$$\begin{aligned} &\frac{dl}{d\beta} = \frac{r^2}{ab} = \frac{p^2}{ab}\frac{1}{(1 - e\cos\beta)^2} = \frac{(1-e^2)^{3/2}}{(1 - e^2\cos\beta)^2} \\ &\frac{dl}{dr} = \pm\frac{r}{a\sqrt{a^2e^2 - (r-a)^2}} \end{aligned} \tag{Q41}$$

It follows from the definition of the eccentric anomaly that $r = a(1 + e\cos\xi)$, and $dr = -ae\sin\xi\, d\xi$, so that:

$$\frac{dl}{d\xi} = \pm\frac{rae\sin\xi\, d\xi}{a\sqrt{a^2e^2 - (r-a)^2}} = 1 + e\cos\xi \qquad l = \xi + e\sin\xi \tag{Q42}$$

and the final choice of the + sign is based on the remark that the average anomaly, the eccentric anomaly and the true anomaly are simultaneously increasing as one of them increased.

The (Q.35), Q.36) are therefore proved, and we have also fond a remarkable formula expressing in a Keplerian motion the mean anomaly in terms of the true anomaly: the first of (Q.42) gives in fact:

$$l = (1 - e^2)^{3/2} \int_0^\beta \frac{d\beta'}{(1 - e\cos\beta')^2} \tag{Q43}$$

which, however, will not be used directly here.

27. *(the gauss' transformation)* Let F be a map transforming a pair (P, Q) of numbers into (P', Q') defined as follows.

Given (P, Q) consider the operations:

(i) solution of the equation for ϱ_2:

$$a\varrho_2 = b + \frac{c-d}{P} + d(1 + \frac{Q}{2r_2^2}) \tag{Q44}$$

(ii)calculation of W via (Q.30), with (P, Q, W) replacing $(\bar{P}, \bar{Q}, \bar{W})$.
(iii)calculation of ϱ_1, ϱ_3 via (Q.31), with (P, Q, W) replacing $(\bar{P}, \bar{Q}, \bar{W})$
(iv)calculation of the elements via (Q.32),(Q.33).
(v)calculation of the parameters $2f_{pq} = \theta_p - \theta_q$ and z_{pq} via (Q.35),(Q.36)
(vi)calculation of P', Q' via:

$$P' = \frac{z_{23}t_{12}}{z_{12}t_{23}}, \qquad Q' = \frac{\kappa t_{12}\kappa t_{23} r_2^2}{r_1 r_3 z_{12} z_{23} \cos f_{12} \cos f_{13} \cos f_{23}} \tag{Q45}$$

Check that, on the basis of the problems 22),23),26), that the analysis developed there can be interpreted as proving that if one sets:

$$P = \frac{n_{12}}{n_{23}}, \qquad Q = (\frac{n_{12} + n_{23}}{n_{12}} - 1)2r_2^3, \tag{Q46}$$

where now n_{pq} and r_2 are the true unknown values of the areas of the triangles S_{pq} and of $|\mathbf{X}_2|$, one has:

$$(P, Q) = F(P, Q) \tag{Q47}$$

at least if one neglects the time aberration, i.e. if one assumes that the time $t_q - t_p$ measured between the observations p and q is the true value of the time interval between the times in which the celestial body occupies the positions p and q, i.e. it can be confused with $t_q - t_p - (\varrho_q - \varrho_p)/c$ (see problem 11)),(*Hint:* check that (Q.44) becomes the first of (Q.15) if (P, Q) are as in (Q.46)).

Write a computer program realizing the map F defined above nd apply it to the computation of (P', Q') in the case of the asteroid Juno using the data given above.

28. *(Gauss' algorithm)* the preceding problem shows that one has to solve (Q.47) as an equation on (P, Q).

We have seen that $(\bar{P}, \bar{Q})$ is a *good* first approximation. It is therefore possible to improve it by some standard methods: the simplest is the *iteration*, another possibility is *Newton's method*. Both were used by Gauss in his boo. And both methods have the drawback that one does not know *a priori* if they will work nor one can easily foretell (if at all possible) an estimate of the time necessary to reach a given precision. Very often they are used empirically and work if one has a *good* approximate solution as a starting point. The methods may otherwise prove an inconclusive or lead to absurd results.

We limit ourselves here to the discussion of the naive iteration method.

Let $\bar{P} = t_{13}/t_{23}$ and $\bar{Q} = \kappa t_{12} \kappa t_{23}$, see problem 24), and define:

$$(P_0, Q_0) = (\bar{P}, \bar{Q}) \qquad (P_k, Q_k) = F(P_{k-1}, Q_{k-1}) \qquad k = 1, 2, \ldots \qquad (Q48)$$

and it is clear that if this makes sense for all k, i.e. if $(P_{k-1}, Q_{k-1}$ is always in the domain of definition of F, then the limit of (P_k, Q_k) as $k \to \infty$ will be, if existing, one solution of the equation and the corresponding data will give the ellipse elements and orbital parameters. Note that the domain of definition of F has not been explicitly defined so far and consists of the set of pairs (P, Q) for which the calculations necessary to evaluate F make sense, i.e. lead to the construction of an ellipse: recall that given three points and a focus there may be no ellipse passing through them; the whole theory can be easily adapted to the case of hyperbolic or parabolic orbits.

In practice one can proceed by starting the iteration from any point (P_0, Q_0). However if this initial point is not close enough to the solution it may happen that (P_k, Q_k) wonders out of the definition domain or has some *strange* asymptotic motion: an undesirable event for our purposes.

The basic difficulty solved by Gauss was to find a method for determining in a rather simple way a first approximation when one knows basically nothing about the asteroid distance; he also devised the above algorithm based on the iteration of a 2-dimensional map, which is remarkably efficient. He showed the power of his method by computing the orbit of the first known asteroid Ceres.

A warning: sometimes the above algorithm may lead to more than one solution as it may be that, even if the original determination of the first approximation for ϱ_2 has a unique acceptable solution, the (Q.47) has more than one fixed points. This could provoke also the unpleasant result that modifications of the algorithm may lead to different final results. Unfortunately it is not easy to develop a general theory of the equation (Q.46) and possible ambiguities have to be solved on an empirical basis.

Use the above scheme to find the elements of the orbit of Juno, on the basis of the data in problem 13).

29. (*correction of time aberrations)* The correction of the time aberrations (problem 11)) can be performed by a small modification of the above iterative method, very easy to implement numerically. Define, if $(\bar{P}, \bar{Q})$ are as in problem 28):

$$\begin{aligned}(P_0, Q_0) =& (\bar{P}, \bar{Q})\\(P_1, Q_1) =& F(P_0, Q_0)\\&\ldots\\(P_{k+1}, Q_{k+1}) =& F_k(P_k, Q_k) \qquad k = 1, 2, \ldots\end{aligned} \tag{Q49}$$

where F_k is obtained from F by replacing t_{pq} in (Q.29) and (Q.45) with: $t_{pq}^{(k)} = t_q - t_p - (\varrho_q^{(k)} - \varrho_p^{(k)})/c$. Check that this leads to the aberration correction. It is simple but it no longer allows to think that the above procedure as an elegant map iteration problem. One could still interpret it as the iteration of a map at the price of increasing the dimension of the space on which the map acts.

Apply the above correction to the elements of Juno.

References for Appendix Q

[1]Gauss, F: *Theory of the motion of heavenly bodies orbiting about the Sun in conical sections*, Dover, N.Y. 1975.
[2]Smart, W.: *Textbook on Spherical Astronomy*, Cambridge U. Press, 1977, Cambridge.
[3]*Explanatory Supplement to the Ephemeris*, jointly prepared by the Nautical Almanac of the U.K. and U.S., London Her Majesty's Stationery Office, 1961 (say).
[4]*The American Ephemeris and Nautical Almanac*, U.S. Government printing Office, Washington, D.C., 20402, 1971 (say).

6.16 S: Definitions and Symbols

Si T est un ensemble, et A une partie de T, on notera φ_A
la fonction charactéristique de A, si cela n'entraine pas de confusion.
(Bourbaki, ch. IX)

$C^\infty(A)$: if $A \subset \mathcal{R}^d$ is an open set: the set of the functions on A continuous, together with their partial derivatives of all orders; shortened often as C^∞ when A is understood.
$C_0^\infty(A)$: if $A \subset \mathcal{R}^d$ is an open set: subset of $C^\infty(A)$ consisting of the functions vanishing outside a closed bounded set contained in A.
$C^{(k)}(A)$: if $A \subset \mathcal{R}^d$ is an open set: it is the set of the functions on A with partial derivatives of order $\leq k$ continuous on A, k being a non-negative integer.
$C^\infty(Q)$: with $Q \subset \mathcal{R}^d$ arbitrary set with dense interior Q_0: set of the functions in $C^\infty(\mathcal{R}^d)$ which vanish outside Q.
$C_0(Q)$: with $Q \subset \mathcal{R}^d$ arbitrary set with dense interior Q_0: set of the functions in $C^\infty(Q_0)$ vanishing outside some closed bounded set contained in Q_0.
$C^{(k)}(Q)$: with $Q \subset \mathcal{R}^d$ arbitrary set with dense interior: defined as $C^\infty(Q)$, considering only the first k derivatives.
$C^\infty(\mathcal{T}^d)$: functions of class C^∞ on the d-dimensional torus $\mathcal{T}^d$ (see Definition 12, p.100, and Definition 13, p.101, §2.21).
$\overline{C}^\infty([0,L])$: functions in $C^\infty([0,L])$ vanishing in 0 and L together with all the even-order derivatives.
$\mathcal{C}^d$: complex d-dimensional space and (or) complex d-dimensional vector space.
$(O;\mathbf{i},\mathbf{j},\mathbf{k})$: orthogonal reference system, O =origin, $\mathbf{i},\mathbf{j},\mathbf{k}$ axes unit vectors.
$\mathcal{R}^d$: : real d-dimensional space and (or) real d-dimensional vector space.
$\mathcal{T}^d$: d-dimensional torus with side 2π (see p.101).
$\mathcal{R}, \mathcal{R}^1$: real line.
$\mathcal{C}, \mathcal{C}^1$: complex plane.
$\mathcal{R}_+$: interval $[0, +\infty)$.
S_t : solution flow for an autonomous differential equation.
$\mathcal{Z}^d$: lattice of the d-tuples of integers.
$\mathcal{Z}, \mathcal{Z}^1$: integer numbers.
$\mathcal{Z}_+$: non-negative integers.
$\boldsymbol{\xi}, \boldsymbol{\eta}, \ldots$: points or vectors in $\mathcal{R}^d$, $\mathcal{C}^d$
$\boldsymbol{\varphi}, \boldsymbol{\psi}, \ldots$: points in $\mathcal{T}^d$.
$(X^{(\alpha)})_{\alpha\in J}$: family of objects $X^{(\alpha)}$ parameterized by in the index set J.
t : real parameter with the interpretation of time.
$\dot{x}$: t-derivative of x.
$\ddot{x}$: second t-derivative of x.
$O(\xi)$: quantity of the order of magnitude of ξ: it means that there is $C > 0$,$\xi_c > 0$ such that $O(\xi) < C|\xi|$ if $|\xi| < \xi_c$. Used when ξ is an "infinitesimal" variable.
$o(\xi)$: quantity infinitesimal of higher order compared to ξ: it means $\lim_{\xi\to 0} |\xi|^{-1} o(\xi) = 0$.
mbe : end-of-proof symbol.
$\mathbf{x} \cdot \mathbf{y}$: scalar product of vectors in $\mathcal{R}^d$.

$\mathbf{x} \wedge \mathbf{y}$	: vector product of two vectors in $\mathcal{R}^3$.
$P - Q$	: vector whose components in a given frame of reference are the differences of the homonymous coordinates of P and Q in the same frame of reference.
$\equiv$	: identity or, often, implicit definition.
$\stackrel{def}{=}$	: implicit definition of l.h.s.by the r.h.s. or viceversa.
$\mathcal{R}$e , $\mathcal{I}$m	: real or imaginary part of a complex number.
$/, \setminus$	: symbols for the set theoretic difference.
∂	: partial derivative or boundary of a set
∂	: gradient operator.

6.17 T: Suggested Books and Complements

A. Suggested Books

Arnold, V.: *Equations Differentielles Ordinaires*, Mir, Moscow, 1976.

Arnold, V.: *Chapitres supplémentaires de la théorie des equations differentielles ordinaires*, Mir, Moscow, 1980.

Arnold, V.: *Méthodes mathématiques de la mécanique classique*, Mir, Moscow, 1978.

Boyer, C.: *History of Mathematics*, Wiley, New York, 1968.

Dreyer, J.: *A History of Astronomy from Thales to Kepler*, Dover, New York, 1953.

Galilei, G.: *Il Saggiatore*, in *Opere*, Edizione Nazionale, Firenze, 1896, vol. VI,

Galilei, G.: *Dialogues Concerning Two New Sciences*, Dover, New York, 1982.

Landau L., Lifschitz, E.: *Mécanique*, Mir, Moscow, 1966.

Lifschitz, E.: *Mécanique des fluides*, Mir, Moscow, 1966.

Levi-Civita, T., Amaldi, V.: *Lezioni di Meccanica Razionale.* Zanichelli, Bologna, 1949.

Mach, E.: *The Science of Mechanics*, Open Court, La SalleLondon, 1942.

Newton, L: *The Mathematical Principles of Natural Philosophy*, Transl. by A. Motte, ed. F. Cajori, Univ. of California Press, Berkeley, 1930.

Sommerfeld, A.: *Atomic Structure and Spectral Lines.* London, 1934.

Truesdell, C.: *Essays in the History of Mechanics*, Springer-Verlag, New York, 1974.

B. Works Developing Themes Introduced in this Book

Arnold, V.: *Small denominators and problems of stability of motion in classical and celestial mechanics*, Russian Mathematical Surveys, **18**(6), 85, (1963).

Arnold, V., Avez, A.: *Ergodic Problems in Classical Mechanics*, Benjamin, New York, 1963.

Gallavotti, G.: *Aspetti della teoria ergodica, qualitativa e statistica del moto*, Quaderni dell' Unione Matematica Italiana, vol. **21**, Pitagora Editrice, Bologna, 1981.

Gnedenko, B.: *The Theory of Probability*, Mir, Moscow, 1973.

Huang, K.: *Statistical Mechanics*, Wiley, New York, 1963.

Kintchin, A.: *Mathematical Foundations of Statistical Mechanics*, Dover, New York, 1957.

Kintchin, A.: *Mathematical Foundations of Information Theory*, Dover, New York, 1963.

Kornfeld, L, Sinai, J. Fomin, S.: *Ergodicescaia Teoria*, Nauka, Moscow, 1980.

Lagrange, L.: *Mécanique analytique*, Paris, 1788.

Lanford, 0.: *Entropy and Equilibrium States in Classical Statistical Mechanics*, Lecture Notes in Physics 20. ed. A. Lenard, Springer-Verlag, Berlin, 1972.

Marsden, J., McCracken, M.: *The Hopf Bifurcation and Its Applications*, Springer-Verlag, Berlin, 1976.

Miranda, C.: *Partial Differential Equations of Elliptic Type*, Springer-Verlag, Berlin, 1970.

Moser, J.: *Lectures on Hamiltonian Systems*, Memoirs of the American Mathematical Society, vol. **81**, 1968.

Moser, J.: *Stable and random motions in dynamical systems*, Annals of mathematical studies, Princeton Univ. Press, Princeton, 1973.

Poincaré, H.: *Les méthodes nouvelles de la mécanique celèste*, vols. 1 and 2. Gauthier-Villars, Paris, 1897.

Reed, M.: *Abstract nonlinear wave equations*, in Lecture Notes in Mathematics **507**. Springer-Verlag, Berlin, 1975.

Ruelle, D.: *Statistical Mechanics. Rigorous Results*, Benjamin, New York, 1969.

References

1. Arnold, V.: *Méthodes mathématiques de la mécanique classique*, Mir, Moscow, 1978, [211,363].
2. Arnold, V.: *A proof of a theorem of A. N. Kolmogorov on the invariance of quasiperiodic motions under small perturbations of the Hamiltonian*, Russian Mathematical Surveys **18**, 9 (1963).[488]
3. Arnold, V.: *Small denominators and problems of stability of motion in classical and celestial mechanics*, Russian Mathematical Surveys, **18**, 85–191, (1963), [493]
4. Arnold, V., Avez, A.: *Ergodic problems of classical mechanics*, Benjamin, New York, 1968. [355,360].
5. Bakouline, P., Kononovitch, E., Monoz, V.: *Astronomie génèrale*, Mir, Moscow, 1975, [546,546,547].
6. Berry, M.: Regular and irregular motions, in AIP Conference proceedings, Vol. 46, ed. S. Jorna. American Institute of Physics, New York, 1978 [302].
7. Bowen, R., Ruelle, D.: *The ergodic theory of axiom A flows*, Inventiones Mathematicae, **29**, 181 (1975) [443].
8. Campanino, M., Epstein, H., Ruelle, D.: *On Feigenbaum's functional equation* $g \circ g(\lambda x) + \lambda g(x) = 0$. Topology, **21**, 125, 1982 [452].
9. Collet, P., Eckmann, J.P.: *Iterated Maps of the Interval as Dynamical Systems*, Birkhauser, Boston, 1980 [452].
10. Courant, R., Hilbert, D.: *Methods of Mathematical Physics*, Vol. II, Interscience, New York, 1962 [333].
11. Chierchia, L.: *Thesis*, Università di Roma, 1981, [514].
12. Chierchia, L., Gallavotti, G.: *Smooth prime integrals for quasi-integrable Hamiltonian systems*, Nuovo Cimento, **67 B**, 277-297 (1982) [490].
13. Deprit, A.: *Free rotations of a rigid body studied in phase space*, American Journal of Physics, **35**, 424 (1967), [318].
14. Fano, G.: *Mathematical Methods of Quantum Mechanics*, McGraw-Hill, New York, 1971, [523,525].
15. Feigenbaum, M.: *Quantitative universality for a class of nonlinear transformations*, Journal of Statistical Physics, **19**, 25 (1978), [451,452].
16. Finzi, B., Udeschini, P.: Esercizi di Mecconica razionale. Tamburini, Milano, 1974, [244].
17. Franceschini, V.: *A Feigenbaum sequence of bifurcations in the Lorenz model*, Journal of Statistical Physics, **22**, 397 (1980), [452].

18. Franceschini, V., Tebaldi, C.: *Sequences of infinite bifurcations and turbulence in a five mode truncation of the Navier-Stokes Equations*, Journal of Statistical Physics, **21**, 707 (1979), [452].
19. Franceschini, V., Tebaldi, C.: *A seven mode truncation of the plane incompressible Navier-Stokes equations*, Journal of Statistical Physics, **25**, 397 (1981), [452].
20. Galilei, G.: *Il Saggiatore*, in *Opere*, Edizione Nazionale, Firenze, 1896, vol. VI, [2]
21. Gallavotti, G.: *Perturbation theory for classical Hamiltonian systems*, in "Scaling and self similarity in Physics", ed. J. Fröhlich, Birkhauser, Boston, 1985, p.359–426, [490].
22. Hadamard, H.: Sur l'iteration et les solutions asymptotiques des equations differentielles, Bulletin Societe Mathématique de France, 29, 224, 1901. [412]
23. Inglese, G.: *Thesis*, Università di Roma, 1981. [513].
24. Levi-Civita, T.: *Opere Matematiche*, Accademia Nazionale dei Lincei, Zanichelli, Bologna, 1956. [486]
25. Khintchin, A.: *Mathematical foundations of Information Theory*, Dover, New York, 1963. [360].
26. Khintchin, A.: *Continued Fractions* Dover, New York, 1964, [96].
27. Kobussen, J.: *Some comments on the lagrangian formalism for systems with general velocity dependent forces.*, Acta Physica Austriaca, **51**, 293 (1979), [138].
28. Landau L., Lifschitz, E.: *Mécanique*, Mir, Moscow, 1966. [48,231,242,310,330, 311]
29. Lanford, 0.: *Bifurcations of periodic solutions into invariant tori. The work of Ruelle and Takens*, in *Lecture Notes in Mathematics*, **322**. Springer-Verlag, Berlin, 1973, [401,412,456].
30. Lanford, 0.: *A computer assisted proof of Feigenbaum's conjecture*, preprint, Berkeley, 1981 [452].
31. Mach, E.: *The Science of Mechanics*, Open Court, La SalleLondon, 1942.[9,211,243,243,243,243]
32. Metcherskij, L.: Recueil de problemes de mecanique rationelle. Mir, Moscow, 1973, [244].
33. Moser, J.: *Lectures on Hamiltonian Systems*, Memoirs of the American Mathematical Society, vol. **81**, 1968, [460,488].
34. Moser, J.: *Stable and random motions in dynamical systems*, Annals of mathematical studies, Princeton Univ. Press, Princeton, 1973, [439,460].
35. Negrini, P., Salvadori, L.: *Attractivity and Hopf bifurcations*, Nonlinear Analysis, **3**, 87, 1978, [407].
36. Nekhorossiev V.: An exponential estimate of the time of stability of nearlyintegrable Hamiltonian systems, Russian Mathematical Surveys, **32**, 1, 1972, [462].
37. Newton, L: *The Mathematical Principles of Natural Philosophy*, Transl. by A. Motte, ed. F. Cajori, Univ. of California Press, Berkeley, 1930.[9,12,300]
38. Poincaré, H.: *Les méthodes nouvelles de la mécanique celèste*, vols. 1 and 2. Gauthier-Villars, Paris, 1897. [18,334]
39. Pöschel, J.: *Integrability of Hamiltonian systems on Cantor sets*, Communications in Pure and Applied Mathematics, **35**, 653 (1982), [490].
40. Rubin, H., Ungar P.: *Motion under a strong constraining force*, Communications on Pure and Applied Mathematics, **10**, 65-87, 1957.
41. Robbins, K.: *Periodic solutions and bifurcation structure at high R in the Lorenz model*, SIAM J. Applied Mathematics, **36**, (1979), [446].

42. Ruelle, D.: *A measure associated with the axiom A attractors*, American Journal of Mathematics, **98**, 619 (1976), [443]..
43. Rüssmann, H.: *Über das Verhalten analytischer Differentialgleichungen in der Nähe einer gleichgewichtslösung*, Mathematische Annalen, **154**, 285 (1964), [363].
44. Rüssmann, H.: Über die Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer gleichgewichtslösung, Mathematische, Annalen, **169**, 55 (1967), [473].
45. Smale, S.: *Differentiable Dynamical Systems*, Bulletin American Mathematical Society, **73**, 747, 1967, [443].
46. Sommerfeld, A.: Atomic . . . [333,333].
47. Tresser, C., Coullet, P.: *Iterations d'endomorphismes et groupe de renormalization*, C. R. Acad. Sci., Paris, **287A**, 577 (1978) [452].
48. Truesdell, C.: *Essays in the History of Mechanics*, Springer-Verlag, New York, 1974.[9,12]
49. Wittaker, E.: A treatise on the analytical dynamics of the rigid body. Cambridge University Press, 1937 [332].

Index